22 February 2013 | $10

Science

HIV & TB
in South Africa

AAAS

cell sciences®

cytokine center

Browse our web site of recombinant proteins, including cytokines, growth factors, chemokines and neurotrophins. Daily shipping and competitive pricing are offered. Bulk quantities of many proteins available. Cell Sciences also carries corresponding antibodies and ELISA kits.

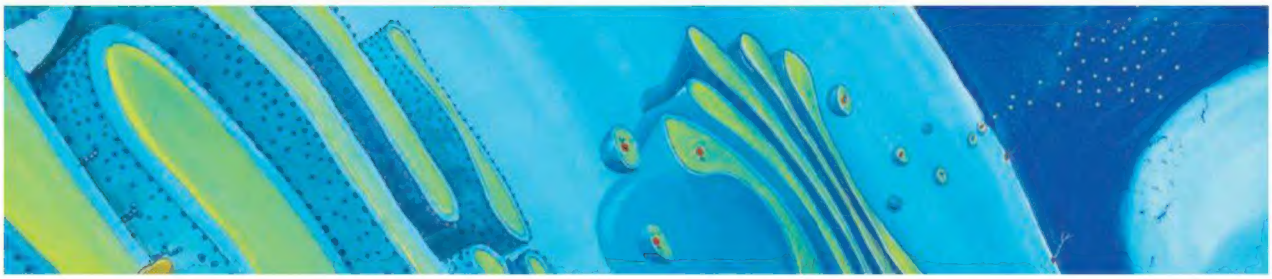
www.cellsciences.com

LIST OF PROTEINS

4-1BBL
4-1BB Receptor
6 Ckine
ACAD8
ACAT2
gAcrp30/Adipolean
Activin A
ACY1
ADAT1
Adiponectin
ADRP
AITRL
Akt1
Alpha-Feto Protein (AFP)
Alpha-Galactosidase A
Angiopoietin-1 (Ang-1)
Angiopoietin-2 (Ang-2)
Angiostatin K1-3
Annexin-V
apo-SAA
Apoliprotein A-1
Apoliprotein E2
Apoliprotein E3
Apoliprotein E4
APRIL
Artemin
ATF2
Aurora A
Aurora B
BAFF
BAFF Receptor
BCA-1 / BLC / CXCL13
BCMA
BD-1
BD-2
BD-3
BDNF
Bivalirudin
BMP-2
BMP-4
BMP-7
BMP-13
sBMPR-1A
Brain Natriuretic Protein
BRAK
Breast Tumor Antigen
C5a
C5L2 Peptide
C-10
C-Reactive Protein
C-Src
Calbindin D-9K
Calbindin D-28K
Calbindin D-29K
Calmodulin
Calcitonin Acetate
Carbonic Anhydrase III
Carcino-embryonic Antigen
Cardiotrophin-1
Caspase-3
Caspase-6
CD4
CD14
CD22
CD40 Ligand / TRAP
CD95 / sFas Ligand
CD105 / Endoglin
CHIPS
CNTF
Collagen
CREB
CTACK/CCL27
CTGF
CTGFL/WISP-2
CTLA-4/Fc
CXCL16
Cytokeratin 8
DEP-1
Desmopressin
Disulfide Oxidoreductase
E-selectin
ECGF
EGF
Elafin/SKALP
EMAP-II
ENA-78
Endostatin
Enteropeptidase
Eotaxin
Eotaxin-2
Eotaxin-3 (TSC)
EPHB2
EPHB4
Eptifbatide
Erk-2
Erythropoietin (EPO)
Exodus-2
Fas Ligand
Fas Receptor
FGF-1 (acidic)
FGF-2 (basic)
FGF-4
FGF-5
FGF-6
FGF-7/ KGF
FGF-8
FGF-9
FGF-10
FGF-16
FGF-17
FGF-18
FGF-19
FGF-20
sFGFR-1 (IIIc) / Fc Chimera
sFGFR-2 (IIIc) / Fc Chimera
sFGFR-3 / Fc Chimera
sFGFR-4 / Fc Chimera
sFlt-1 (native)
sFlt-1 (D3)
sFlt-1 (D4)
sFlt-1 (D5)
sFlt-1 (D7)
Flt3-Ligand
sFlt-4
sFlt-4/ Fc Chimera
Follistatin
FSH
Fractalkine/ CX3C
G-CSF
α-Galactosidase A
Galectin-1
Galectin-3
Gastrointestinal CA
GCP-2
GDF-3
GDF-9
GDF-11
GDNF
GLP-1
Glucagon
Goserelin
GM-CSF
GPBB
GROα
GROβ
GROγ
GRO/MGSA
Growth Hormone
Growth Hormone BP
GST-p21/WAF-1
HB-EGF
HCC-1
HGF
Histdyl-tRNA synthetase
Histrelin
HRG1-β1
I-309
I-TAC
IFN-α
IFN-α A
IFN-α 2a
IFN-α 2b
IFN-β
IFN-γ
IFN-Omega
IGF-I
IGF-II
proIGF-II
IGFBP-1
IGFBP-2
IGFBP-3
IGFBP-4
IGFBP-4
IGFBP-5
IGFBP-6
IGFBP-7
IL-1α
IL-1β
IL-2
IL-3
IL-4
sIL-4 Receptor
IL-5
IL-6
sIL-6 Receptor
IL-7
IL-8 (72 a.a.)
IL-8 (77 a.a.)
IL-9
IL-10
IL-11
IL-12
IL-13
IL-13 analog
IL-15
IL-16 (121 a.a.)
IL-16 (130 a.a.)
IL-17
IL-17B
IL-17D
IL-17E
IL-17F
IL-19
IL-20
IL-22
IL-31
Insulin
IP-10
JE
JNK2a1
JNK2a2
KC / CXCL1
KGF
L-asparaginase
LAG-1
LALF Peptide
LAR-PTP
LC-1
LBP
LD-78β
LDH
LEC/NCC-4
Leptin
LIGHT
LIX
LKM
LL-37
Lymphotactin
sLYVE-1
M-CSF
MCP-1 (MCAF)
MCP-2
MCP-3
MCP-4
MCP-5
MDC (67 a.a.)
MDC (69 a.a.)
MDH
MEC
Mek-1
MIA
Midkine
MIG / CXCL9
MIP-1α / CCL3
MIP-1β / CCL4
MIP-3 / CCL23
MIP-3α / CCL20
MIP-3β / CCL19
MIP-4 (PARC) / CCL18
MIP-5 / CCL15
MMP-3
MMP-7
MMP-13
Myostatin
Nanog
NAP-2
Neurturin
NFAT-1
beta-NGF
NOGGIN
NOV
NP-1
NT-1/BCSF-3
NT-3
NT-4
Ocreotide
Oncostatin M
Osteoprotegerin (OPG)
OTOR
Oxytocin
p38-α
Parathyroid Hormone
PDGF-AA
PDGF-AB
PDGF-BB
Persephin
PF-4
PIGF-1
PIGF-2
PKA α-subunit
PKC-α
PKC-γ
Pleiotrophin
PLGF-1
Polymyxin B (PMB)
PRAS40
PRL-1
PRL-2
PRL-3
Prokineticin-2
Prolactin
Protirelin
PTHrP
PTP1B
PTP-IA2
PTP-MEG2
PTP-PEST
sRANK
sRANKL
RANTES
RELM-α
RELM-β
Resistin
RPTPβ
RPTPγ
RPTPμ
SCF
SCGF-α
SCGF-β
SDF-1α
SDF-1β
Secretin
SF20
SHP-2
STAT1
c-Src
TACI
TARC
TC-PTP
TECK
TFF2
TGF-α
TGF-β1
TGF-β2
TGF-β3
Thymosin α1
sTIE-1/Fc Chimera
sTIE-2/Fc Chimera
TL-1A
TNF-α
TNF-β
sTNFR1
sTNFR2
TPO
TRAIL/Apo2L
sTRAIL R-1 (DR4)
sTRAIL R-2 (DR5)
TSH
TSLP
TWEAK
TWEAK Receptor
Urokinase
VEGF121
VEGF145
VEGF165
VEGF-C
VEGF-C I525
EG-VEGF
VEGF-E
HB-VEGF-E
sVEGFR-1
sVEGFR-2
sVEGFR-3
WISP-1
WISP-2
WISP-3
WNT-1

CELL SCIENCES INC • 480 NEPONSET STREET, BUILDING 12A, CANTON, MA 02021 • INFO@CELLSCIENCES.COM

TOLL FREE: (888) 769-1246 • TEL: (781) 828-0610 • FAX: (781) 828-0542 • WEB: WWW.CELLSCIENCES.COM



R&D Systems Cytokines & Growth Factors

Setting the Standard for the Life Science Industry for over 25 Years

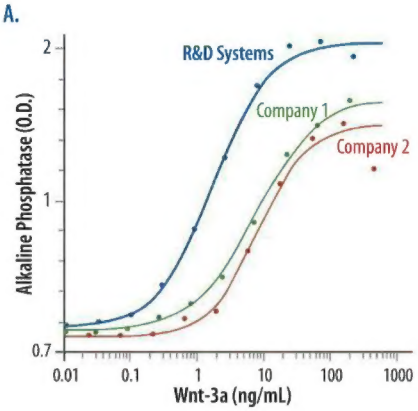
✓ R&D Systems currently offers more than 3,000 bioactive recombinant proteins

✓ We are the most referenced manufacturer of recombinant & natural proteins

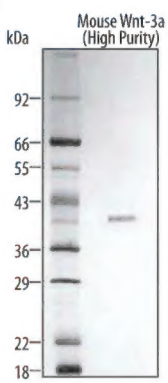
✓ Our scientists are dedicated to developing the highest quality products & new processes that meet the needs of researchers

High purity proteins, traditional & cell culture grade • Animal-free proteins • GMP proteins

A.



B.



A. Activity Comparison Data for Recombinant Mouse Wnt-3a. The MC3T3-E1 mouse preosteoblast cell line was treated with increasing concentrations of R&D Systems Recombinant Mouse Wnt-3a (Catalog # 1324-WN) or with mouse Wnt-3a from two other companies. The R&D Systems protein shows greater than a four-fold increase in activity compared to the other commercially available proteins.

B. SDS-PAGE Analysis of High Purity Recombinant Mouse Wnt-3a. High Purity Recombinant Mouse Wnt-3a (Catalog # 1324-WNP; 1 µg/lane) was loaded on a 12% SDS-PAGE gel under reducing conditions and visualized by silver staining.

www.RnDSystems.com/CytokinesGrowthFactors

R&D Systems Tools for Cell Biology Research™

For research use only. Not for use in diagnostic procedures.

R&D SYSTEMS®

Explore the fascinating new world of QIAGEN®

Virtual showroom and application labs

Enjoy discovering innovative sample and assay technologies at new QIAGEN.com

- Experience new dimensions
- Discover fascinating innovations
- Relevant for your personal applications

Scan the code with your iPad,
or go to www.qiagen.com/qiagen-experience.

Experience

QIAGEN®

Sample & Assay Technologies

page 893

page 915

CONTENTS continued >>

COVER

With her mother looking on, a woman in Hlabisa, South Africa, receives an HIV test as part of a survey by the Africa Centre for Health and Population Studies. In this issue, two Reports (pages 961 and 966) show that treating HIV-infected people reduces the risk of virus spread and increases life expectancy. A special News section (page 898) describes South Africa's fight against the co-epidemics of HIV and tuberculosis.

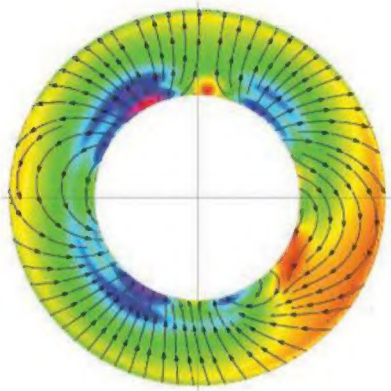
Photo: Malcolm Linton

DEPARTMENTS

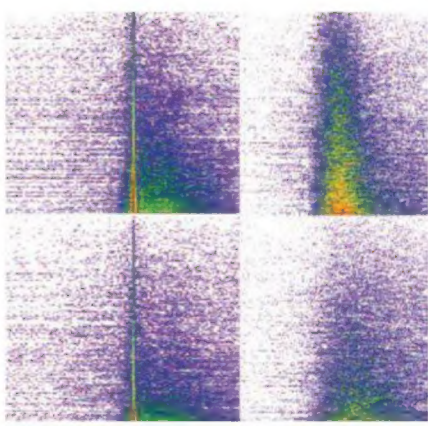
REVIEW

RESEARCH ARTICLE

REPORTS



page 928



page 950

SCIENCE (ISSN 0036-8075) is published weekly on Friday, except the last week in December, by the American Association for the Advancement of Science, 1200 New York Avenue, NW, Washington, DC 20005. Periodicals Mail postage (publication No. 484460) paid at Washington, DC, and additional mailing offices. Domestic individual membership and subscription (51 issues): $149 ($74 allocated to subscription). Domestic institutional subscription (51 issues): $990; Foreign postage extra: Mexico, Caribbean (surface mail) $55; other countries (air assist delivery) $85. First class, airmail, student, and emeritus rates on request. Canadian rates with GST available upon request, GST #1254 88122. Publications Mail Agreement Number 1069624. **Printed in the U.S.A.**

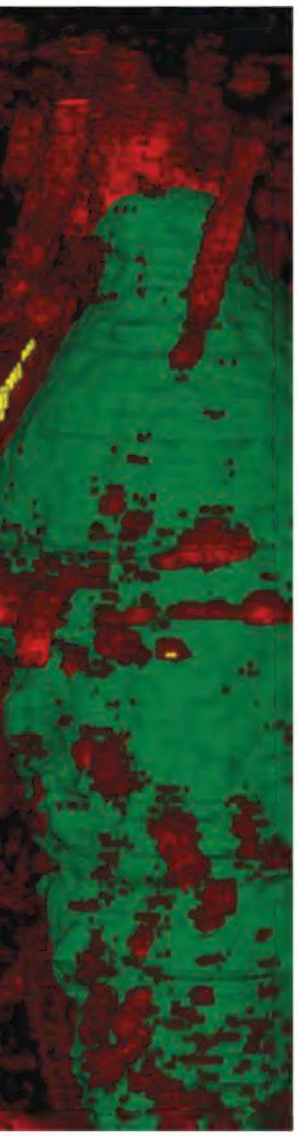
Change of address: Allow 4 weeks, giving old and new addresses and 8-digit account number. **Postmaster:** Send change of address to AAAS, P.O. Box 96178, Washington, DC 20090–6178. **Single-copy sales:** $10.00 current issue, $15.00 back issue prepaid includes surface postage; bulk rates on request. *Science* is indexed in the *Reader's Guide to Periodical Literature* and in several specialized indexes.

American Association for Cancer Research

SCIENTIFIC CONFERENCES 2013:

Presenting the most significant research on cancer etiology, prevention, diagnosis, and treatment

www.aacr.org/Conferences2013



AACR Annual Meeting 2013
Chairperson: José Baselga
April 6-10, 2013 • Washington, DC

Synthetic Lethal Approaches to Cancer Vulnerabilities
Co-Chairpersons: William C. Hahn, Sebastian Nijman, and Louis M. Staudt
May 17-20, 2013 • Bellevue, WA

Chromatin and Epigenetics in Cancer
Co-Chairpersons: Suzanne J. Baker, Charles W.M. Roberts, and Gerald R. Crabtree
June 19-22, 2013 • Atlanta, GA

Frontiers in Basic Cancer Research
Chairperson: Scott W. Lowe
Co-Chairpersons: Joan S. Brugge, Hans Clevers, Carol L. Prives, and Davide Ruggero
September 18-22, 2013 • National Harbor, MD

Advances in Ovarian Cancer Research: From Concept to Clinic
Co-Chairpersons: David G. Huntsman, Douglas A. Levine, and Sandra Orsulic
September 18-21, 2013 • Miami, FL

Advances in Breast Cancer Research
Co-Chairpersons: Carlos L. Arteaga, Jeffrey M. Rosen, Jane E. Visvader, and Douglas Yee
October 3-6, 2013 • San Diego, CA

AACR-NCI-EORTC International Conference on Molecular Targets and Cancer Therapeutics
Co-Chairpersons: Jeffrey A. Engelman, Lee J. Helman, and Sabine Tejpar
October 19-23, 2013 • Boston, MA

Twelfth Annual International Conference on Frontiers in Cancer Prevention Research
Chairperson: Paul J. Limburg
October 27-30, 2013 • National Harbor, MD

Pediatric Cancer at the Crossroads: Translating Discovery into Improved Outcomes
Co-Chairpersons: John M. Maris, Stella M. Davies, James R. Downing, Lee J. Helman, and Michael B. Kastan
November 3-6, 2013 • San Diego, CA

The Translational Impact of Model Organisms in Cancer
Co-Chairpersons: Cory Abate-Shen, A. Thomas Look, and Terry A. Van Dyke
November 5-8, 2013 • San Diego, CA

CTRC-AACR San Antonio Breast Cancer Symposium
Co-Directors: Carlos L. Arteaga, C. Kent Osborne, and Peter M. Ravdin
December 10-14, 2013 • San Antonio, TX

Sixth AACR Conference on The Science of Cancer Health Disparitites in Racial/Ethnic Minorities and the Medically Underserved
December 2013 • Location to be determined

AACR EDUCATIONAL WORKSHOPS

Accelerating Anticancer Agent Development and Validation
Co-Chairpersons: H. Kim Lyerly and Richard Pazdur
May 8-10, 2013
Bethesda, MD

NEW! Integrative Molecular Epidemiology
Director: Thomas A. Sellers; Co-Directors: Peter L. Kraft and Margaret R. Spitz
July 15-20, 2013
Boston, MA

Molecular Biology in Clinical Oncology
Co-Directors: William G. Kaelin Jr., Mark Geraci, and Suzanne Topalian
July 21-28, 2013
Snowmass, CO

ASCO/AACR Methods in Clinical Cancer Research
Co-Directors: Jamie H. von Roenn, Neal J. Meropol, and Mithat Gönen
July 27-August 2, 2013
Vail, CO

Translational Research for Basic Scientists
Co-Directors: Tom Curran, George D. Demetri, and Pasi A. Jänne
November 10-15, 2013
Boston, MA

www.aacr.org/Workshops2013

In Search of Unparticles

The standard model of particle physics, which describes the basic building blocks of the universe and the interactions among them, is incomplete. Numerous theoretical extensions have been proposed, some of which predict long-range, spin-spin interactions. To test whether such interactions exist, a laboratory spin source is normally used. **Hunter *et al.*** (p. 928) used Earth as a polarized spin source and looked for these interactions by changing the geographical position and the orientation of the measurement apparatus. The polarized spins mainly come from the electrons in iron-containing minerals of Earth's mantle, which align in Earth's magnetic field. The large numbers of such polarized electrons allowed the extraction of upper bounds on some of the exotic spin-spin interactions far lower than those obtained in the laboratory.

Cost-Benefit of ART

In the battle to control HIV, mass antiretroviral treatment (ART) costs $500 to $900 per person per year. **Bor *et al.*** (p. 961) calculated the impact of intensifying ART on the life expectancy of people living in rural KwaZulu Natal. The dates of death were collected from a population of about 100,000 people during 2000–2011: Four years before and 8 years after the scaling up of ART. Life expectancy of adults increased by more than 11 years after ART was expanded, and the economic value of the lifetimes gained were calculated to far exceed the cost of treatment. **Tanser *et al.*** (p. 966) followed nearly 17,000 HIV-uninfected individuals in KwaZulu-Natal over an 8-year period. Holding other HIV risk factors constant, individual HIV acquisition risk declined significantly with increasing ART coverage of HIV-infected people.

Mother Knows Best

Fruit flies feed on decomposing fruit, a food source rich in alcohol, and thus have evolved a high alcohol tolerance not shared by many other species. High levels of ingested alcohol protect fly larvae against parasitic wasp larvae. **Kacsoh *et al.*** (p. 947) show that this therapeutic use of alcohol to combat a parasite has an intergenerational component. When fly mothers were permitted to see parasitic wasps, they preferentially laid their eggs on substrates containing high alcohol levels as a way to medicate their offspring against potential infection.

CREDITS (TOP TO BOTTOM): KULMALA *ET AL.*; ALEX WILD/ALEXANDERWILD.COM

Aerosol Formation

Most atmospheric aerosol particles result from a growth process that begins with atmospheric molecules and clusters, progressing to larger and larger sizes as they acquire other molecules, clusters, and particles. The initial steps of this process involve very small entities—with diameters of less than 2 nanometers—which have been difficult to observe. Kulmala *et al.* (p. 943; see the Perspective by Andreae) developed a sensitive observational protocol that allows these tiny seeds to be detected and counted, and they mapped out the process of aerosol formation in detail.

Water Flowing Underground

In addition to serving as an out-of-sight, yet much-needed water source, groundwater influences ecosystems on land—especially when the groundwater depth is shallow. **Fan *et al.*** (p. 940) used government archives and published studies to construct a global map of groundwater depth based on over 1,000,000 direct well measurements. A groundwater model was then used to construct a continuous global map of groundwater depth. The findings reveal the global influence of sea level and climate on groundwater depths across several regions and ecosystems.

Stealth Delivery

Delivery of therapeutics and imaging agents is hampered by the ability of the innate immune system to recognize and clear foreign particles. "Self" cells are protected from phagocytic clearance by the membrane protein CD47 that interacts with signal regulatory protein-α (SIRPα) on macrophages. Taking advantage of this protective strategy, **Rodriguez *et al.*** (p. 971) labeled nanoparticles with computationally designed minimal human-CD47–based peptide (hCD47). When injected into a strain of mice in which macrophages expressed a SIRPα that cross-reacts with hCD47, the peptide prevented clearance of nanoparticles and enhanced drug delivery to tumors.

Assessing Singletons

The one-child policy introduced by the government of China in 1979 increased the proportion of urban families with an only child; later referred to as "little emperors" in media reports. In 2010, **Cameron *et al.*** (p. 953, published online 10 January) recruited approximately 400 residents of Beijing who had been born either before the implementation of the policy (1975 and 1978) or after (1980 and 1983). Using economic games to measure trust, risk, and willingness to compete, they found that the post-1979 cohorts were less trusting and less willing to compete and also more risk averse.

Light On Clusters

From schools of fish to bacterial colonies, large-scale phenomena—such as swarming or pattern formation—are ubiquitous. In such systems, there is a continuing question as to the relative importance of "intelligence" (biology) of the agents versus purely physical effects. Working with synthetic colloids, **Palacci *et al.*** (p. 936) show that self-organized clustering can be switched on and off to form crystals that dissolve when the light source is turned off. The particles consist of a hematite cube partly encapsulated in a polymeric sphere that is able to catalyze chemical reactions when exposed to light. The self-assembly behavior results from a combination of propelling forces, osmotic effects, and coupling between colloidal and tracer particles.

Louise O. Fresco is a University Professor at the University of Amsterdam, the Netherlands, and an Honorary Professor at the Wageningen University and Research Center, the Netherlands. From 1996 to 2006, she served as Director of Research and Assistant Director General for Agriculture at the Food and Agriculture Organization of the United Nations. E-mail: L.O.Fresco@uva.nl.

The GMO Stalemate in Europe

IN SEPTEMBER 2012, EUROPE WAS SHOCKED BY A PUBLICATION, FROM CAEN UNIVERSITY IN FRANCE, claiming that rats fed for 2 years with transgenic herbicide-resistant corn suffered from tumors. Even though the results have been criticized as flawed,* this research continues to be hailed as a confirmation that genetically modified organisms (GMOs) are intrinsically dangerous.

The European Union (EU) differs from most of the world in its strong opposition to the use of genetic modification in agriculture. This position has worsened over the past 15 years. Field trials of new GM varieties have declined since the late 1990s. Nearly the entire EU commercial acreage of 100,000 ha consists of Bt corn, altered to express a toxin from *Bacillus thuringiensis* that is poisonous to insect pests; no other GM crop is allowed, apart from a high-starch potato. The corn is grown mostly in Spain, the only European country in the top 20 GM crop–growing countries worldwide. Once the European Food Safety Authority has produced a positive "final opinion" concerning the suitability of a new GM crop, final authorization must come from the European Commission (EC) and the member states that vote on approval. Over a dozen GM crops are stuck somewhere in this pipeline, some stalled for years, either because of the absence of support from a majority of member states or a failure of the EC to submit the case to a vote. Attempts to break the deadlock have included seeking an agreement that would allow an individual member state to block the cultivation of a particular GM crop on its own territory, based on safety issues, while allowing other EU nations to make a decision about growing it. Unfortunately, such efforts to ease acceptance of genetic modification have failed.

Respected independent institutions in Europe have provided evidence that GM crops can contribute to sustainable food production, especially when bred for insect and disease resistance, and that they do not carry risks beyond those of conventional varieties.† In 2011, the EC stated that the authorization procedure is dominated by preconceived ideas that prevent a fair revision of procedures to evaluate, approve, and control GMOs. However, in reaction to the flawed Caen study, the EC has opted for further delay, seeking more research on the long-term effects of GM feed. Yet 39 GM crops are currently allowed into the EU as food or feed, with many new requests expected. Europeans and their livestock are already consuming GM foods on a substantial scale.

Europe's lack of trust in GMOs reflects a wider distrust of science. Similar attitudes prevail concerning shale gas and nuclear power. The irony is that the generations who have benefited most from scientific progress are now the most suspicious of science. Europeans tend to romanticize the pre-modern past, unaware of the suffering and food scarcity associated with low crop yields. This European distrust of science affects R&D investments and may have harmful effects elsewhere. In Africa, European donors and nongovernment organizations (NGOs) unnecessarily delay the introduction of disease-resistant GM plants, such as the cassava needed to counteract the growing famine caused by brown streak virus.‡

A change in European attitudes will not arise quickly. Nevertheless, this year's negotiations for the renewal of the EU Common Agricultural Policy for 2014–2020 may provide an opportunity, if the revision of subsidies is coupled with support for innovations, including GMOs that promote sustainable agriculture. Only political courage, as shown last year by the British government's request for the EU to make it easier to grow GMOs, can break the ideological stalemate between NGOs, producers, consumers, and scientists. **– Louise O. Fresco**

10.1126/science.1236010

*European Food Safety Authority, *EFSA J.* **10**, 2986 (2012). †Swiss National Research Program, *Benefits and Risks of the Deliberate Release of Genetically Modified Plants* (2012), www.nfp59.ch/e_index.cfm; The Royal Society, *Genetically Modified Plants for Food Use and Human Health—An Update* (2002), http://royalsociety.org/uploadedFiles/Royal_Society_Content/policy/publications/2002/9960.pdf; A. C. Franke *et al.*, *Sustainability of Current GM Crop Cultivation* (2012), http://edepot.wur.nl/166665. ‡H. Vanderschuren *et al.*, *PLoS ONE* **7**, e45277 (2012).

EDITED BY KRISTEN MUELLER AND MARIA CRUZ

CLIMATE SCIENCE

Modern Warming by Proxy

Records constructed by compiling direct measurement of surface air temperatures made at thousands of locations worldwide over the past century or so show clearly and repeatedly that climate is warming globally. Surface air temperatures for periods before direct measurements were made systematically and broadly enough to be useful on global scales are inferred from sources such as tree rings, corals, ice cores, speleothems, and marine sediments, which provide proxies for temperature. What do proxy records say about modern global warming, though? Anderson *et al.* present an independent record of climate warming over the past 130 years, assembled from a collection of 170 temperature-sensitive paleo proxies, which show the same warming trend as the instrumental record over the same time, including the more detailed picture of warming from around 1910 to the 1940s, the pause in temperature rise until the 1960s, and a continuation of warming from the 1960s to the present. This work validates both the modern instrumental record of temperature rise and the use of these proxies to reconstruct the temperatures of the past. — HJS

Geophys. Res. Lett. **40**, 189 (2013).

ECONOMICS

Risky Fishing

Fishing is a risky way to make a living, not only in terms of physical injury and death but also financially. Although farmers often have access to crop insurance, price supports, and futures markets to mitigate risk, such tools are unavailable to commercial fishers. One approach common to both farmers and fishers, however, is to diversify their harvests. But fishers face constraints on their portfolio due to limited licenses and catch quotas imposed by fisheries managers. To better understand this phenomenon, Kasperski and Holland analyzed harvest and revenue data for over 30,000 commercial fishing vessels operated off the U.S. West Coast and Alaska from 1981 to 2010. Maximum income variability was observed when harvests were only modestly diverse, roughly a 90-10% split between the two fisheries. Although income was slightly more stable when harvests were not at all diverse (100-0%), income stability was more pronounced as harvests were spread 50-50% across two fisheries, or even more stable when spread 50-25-25% across three fisheries. Fishery restrictions may be necessary to ensure sustainability and efficiency; however, they do have important economic effects on fishers. — BW

Proc. Natl. Acad. Sci. U.S.A. **110**, 2076 (2013).

NEUROSCIENCE

Tuning Brain Blood Flow

Although the brain is insulated from many systemic effects on blood flow, delicately regulated changes in blood flow bring resources to sites of high neuronal activity. Astrocytes are in contact with neurons, and astrocyte endfeet also touch the small blood vessels that penetrate the brain. Cerebral arteries dilate in response to glutamate and D-serine signaling through NMDA receptors. LeMaistre Stobart *et al.* have studied the signaling pathways by which astrocytes regulate the blood flow in arterioles in mouse brain slices and found that astrocytes can be a source of endogenous D-serine. D-serine found in the endfeet of astrocytes was released in response to the stimulation of glutamatergic neurotransmission. The subsequent vascular dilation required nitric oxide synthase and nitroic oxide signaling from intact vascular endothelium. These results suggest that astrocytes activated by neighboring neural activity call in additional blood flow by signaling for the local dilation of blood vessels. — PJH

Proc. Natl. Acad. Sci. U.S.A. **110**, 10.1073/pnas.1215929110 (2013).

DEVELOPMENT

Old Stem Cells Rejuvenated

Hematopoietic stem cells (HSCs) in the bone marrow give rise to all cells of the blood, but with age, this capacity decreases. Sirtuin 3 (SIRT3) is a deactylase expressed in mitochondria that can control the production of damaging reactive oxygen species (ROS) by, for example, improving the antioxidative activity of other enzymes. Brown *et al.* observed that SIRT3 is highly expressed in the HSCs of young mice but not in older mice. In mice genetically engineered to lack SIRT3, the number of HSCs declined in older animals and, consequently, reduced the number of differentiated blood cells. In contrast, young mice lacking SIRT3 showed no such effects. In bone marrow transplant assays, only HSCs from young wild-type mice or young SIRT3-deficient mice could reconstitute blood in recipient mice. SIRT3-deficient HSCs from old mice also were susceptible to oxidative stress in vitro and showed reduced survival in vivo. The overexpression of SIRT3 in older SIRT3-deficient HSCs reduced ROS production and boosted the production of blood cells in bone marrow transplantation assays. Thus, the reduced expression of SIRT3 with age probably disrupts mitochondrial

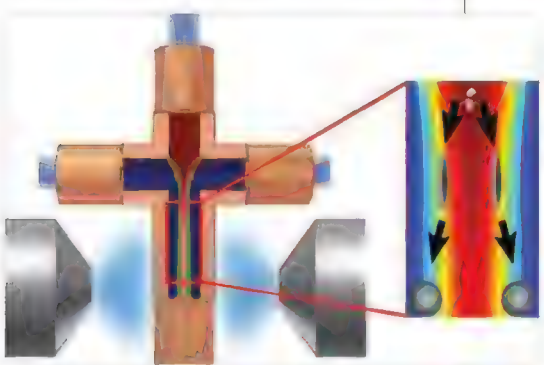
homeostasis and increases oxidative stress, thereby compromising stem cell renewal and function. Age-related stem cell degeneration may be reversed by manipulating mitochondrial homeostasis. — LC

Cell Rep. 10.1016/j.celrep.2013.01.005 (2013).

APPLIED PHYSICS

A Quick Freeze

Microfluidics, in which small volumes of liquids are manipulated inside narrow channels, has proven to be a useful tool for the controlled fabrication of a range of materials. For example, through the delicate mixing of two fluid streams, it is possible to produce uniform oil-in-water emulsion droplets. Because the flow is laminar, it is also possible to create a long mixing layer at the interface between two miscible fluids. Cryogenic freezing has been used to capture or preserve fragile biological structures while minimizing ice crystal formation for subsequent analysis with an electron microscope. Jahn *et al.* extended this thinking to devise a technique for rapidly freezing a capillary stream as it exits the microfluidic device using jets of liquid propane. They studied the formation of nanometer-sized lipid membrane structures such as unilamellar vesicles, which are of interest because of the use of liposomes in cosmetics, pharmaceuticals, and food production. Through this process, they were able to capture both fully formed structures and their intermediaries. For the assembly of phosphocholines at the water and isopropyl alcohol interfaces, they observed disklike intermediate assemblies, consistent with current theories on the assembly process. — MSL



CREDIT: A. JAHN *ET AL.*, *LANGMUIR* **29**, 1717 (4 JANUARY 2013) © 2013 AMERICAN CHEMICAL SOCIETY

Langmuir **29**, 1717 (2013).

PHYSICS

Voltage-Controlled Magnetism

Many components of modern devices are based on magnetic effects; magnetic read heads, for example, use the phenomenon known as exchange bias, in which two materials that order magnetically are interfaced, causing the magnetic properties of the ferromagnetic layer (in which spins align parallel to each other) to be affected by the antiferromagnetic layer (in which neighboring spins prefer to be antiparallel). In devices, the control of magnetic effects with electric fields is desirable. Building on earlier work that showed a change in the size of the exchange bias effect, Wu *et al.* demonstrate a reversible switching of the sign of the exchange bias after the application of a voltage pulse. Key to this result is the use of a multiferroic material for the antiferromagnetic layer, in which switching the sign of the voltage pulse causes spins to reverse their direction and changes the distance between the magnetic ions in the antiferromagnetic and ferromagnetic layers; this, in turn, causes the exchange bias to switch. The switching is achieved at constant temperature (without cooling or heating cycles) and without applying additional magnetic fields; the mechanism the authors offer as an explanation may aid in the understanding of the exchange bias effect. — JS

Phys. Rev. Lett. **110,** 067202 (2013).

IMMUNOLOGY

Regulating TLR Traffic Flow

Sensing pathogen-derived nucleic acids is important for mounting antipathogen immunity. Such nucleic acid–sensing receptors are expressed intracellularly so that they do not inappropriately respond to self nucleic acids and cause autoimmunity. The specific mechanisms that regulate how these receptors, which include Toll-like receptors (TLRs) 3, 7, 8, and 9, reach their specific locations within cells, however, are not well understood. Lee *et al.* use cell biology techniques and mutational analysis to demonstrate that the transmembrane endoplasmic reticulum (ER) protein UNC93B1 is important for this process. UNC93B1 assists in the packaging of at least six TLRs in the ER and then remains associated with these receptors as they are processed through the Golgi apparatus. How TLRs go from the Golgi to their final destination, however, appears to be regulated by distinct mechanisms depending on the receptor. UNC93B1-dependent recruitment of AP-2 was required for TLR9 trafficking to endosomes; however, AP-2 was not required by other receptors. Thus, a variety of mechanisms may be employed to get TLRs from the Golgi to the right locations within the cell. — KLM

eLife 10.7554/eLife.00291 (2013).

www.sciencemag.org

Science

1200 New York Avenue, NW
Washington, DC 20005
Editorial: 202-326-6550, FAX 202-289-7562
News: 202-326-6591, FAX 202-371-9227
Bateman House, 82-88 Hills Road
Cambridge, UK CB2 1LQ
+44 (0) 1223 326500, FAX +44 (0) 1223 326501

Subscription Services For change of address, missing issues, new orders and renewals, and payment questions: 866-434-AAAS (2227) or 202-326-6417, FAX 202-842-1065. Mailing addresses: AAAS, P.O. Box 96178, Washington, DC 20090-6178 or AAAS Member Services, 1200 New York Avenue, NW, Washington, DC 20005

Institutional Site Licenses please call 202-326-6755 for any questions or information

Reprints: Author Inquiries 800-635-7181
Commercial Inquiries 803-359-4578

Permissions 202-326-7074, FAX 202-682-0816

Member Benefits AAAS Travels: Betchart Expeditions 800-252-4910; Apple Store www.store.apple.com/us/go/eppstore/aaas; NASA Federal, 1-888-NASA-FCU (1-888-627-2328) or www.nasafcu.com; Cold Spring Harbor Laboratory Press Publications www.cshlpress.com/affiliates/aaas.htm; GEICO Auto Insurance www.geico.com/landingpage/go51.htm?logo=17624; Hertz 800-654-2200 CDP#343457; Office Depot https://bsd.officedepot.com/portalLogin.do; Seabury & Smith Life Insurance 800-424-9883; Subaru VIP Program 202-326-6417; VIP Moving Services www.vipmayflower.com/domestic/index.html; Other Benefits: AAAS Member Services 202-326-6417 or www.aaasmember.org.

science_editors@aaas.org	(for general editorial queries)
science_letters@aaas.org	(for queries about letters)
science_reviews@aaas.org	(for returning manuscript reviews)
science_bookrevs@aaas.org	(for book review queries)

Published by the American Association for the Advancement of Science (AAAS), *Science* serves its readers as a forum for the presentation and discussion of important issues related to the advancement of science, including the presentation of minority or conflicting points of view, rather than by publishing only material on which a consensus has been reached. Accordingly, all articles published in *Science*—including editorials, news and comment, and book reviews—are signed and reflect the individual views of the authors and not official points of view adopted by AAAS or the institutions with which the authors are affiliated.

AAAS was founded in 1848 and incorporated in 1874. Its mission is to advance science, engineering, and innovation throughout the world for the benefit of all people. The goals of the association are to: enhance communication among scientists, engineers, and the public; promote and defend the integrity of science and its use; strengthen support for the science and technology enterprise; provide a voice for science on societal issues; promote the responsible use of science in public policy; strengthen and diversify the science and technology workforce; foster education in science and technology for everyone; increase public engagement with science and technology; and advance international cooperation in science.

Information for Authors
See pages 716 and 717 of the 8 February 2013 issue or access www.sciencemag.org/about/authors

EDITOR-IN-CHIEF **Bruce Alberts**
EXECUTIVE EDITOR **Monica M. Bradford**
NEWS EDITOR **Colin Norman**
MANAGING EDITOR, RESEARCH JOURNALS **Katrina L. Kelner**
DEPUTY EDITORS **R. Brooks Hanson, Barbara R. Jasny, Andrew M. Sugden, Valda J. Vinson**

Editorial senior editors/commentary Lisa D. Chong, Brad Wible; **senior editors** Gilbert J. Chin, Pamela J. Hines, Paula A. Kiberstis (Boston), Marc S. Lavine (Toronto), Beverly A. Purnell, L. Bryan Ray, Guy Riddihough, H. Jesse Smith, Phillip D. Szuromi (Tennessee), Jake S. Yeston, Laura M. Zahn (San Diego); **associate editors** Melissa R. McCartney (Education Projects), Kristen L. Mueller, Jelena Stajic, Sacha Vignieri (Oregon), Nicholas S. Wigginton; **book review editor** Sherman J. Suter; **associate letters editor** Jennifer Sills; **editorial manager** Cara Tate; **senior copy editors** Jeffrey E. Cook, Cynthia Howe, Harry Jach, Lauren Kmec, Barbara P. Ordway, Trista Wagoner; **copy editor** Chris Filiatreau; **senior editorial coordinators** Carolyn Kyle, Beverly Shields; **editorial coordinators** Joi S. Granger, Anita Wynn; **publications assistants** Ramatoulaye Diop, Aneera Dobbins, Jeffrey Hearn, Lisa Johnson, Dona Mathieu, Le-Toya Mayne Flood, Shannon McMahon, Scott Miller, Jerry Richardson, Teresa R. Sakon, Brian White; **editorial assistant** Patricia M. Moore; **executive editorial assistant** Yolanda O'Bannon (San Francisco); **executive assistant** Alison Crawford; **administrative support** Maryrose Madrid

Editorial director, web and new media Stewart Wills; **senior web editor** Sarah Crespi; **web editor** Kerry Klein; **web development manager** Martyn Green; **web developer** Corinna Cohn

News deputy news editors Robert Coontz, Elizabeth Culotta, David Grimm (Online), Eliot Marshall, Jeffrey Mervis, Leslie Roberts, Richard Stone, John Travis; **contributing editor** Polly Shulman; **news writers** Yudhijit Bhattacharjee, Adrian Cho, Jennifer Couzin-Frankel, Carolyn Gramling, Jocelyn Kaiser, Richard A. Kerr, David Malakoff, Elizabeth Pennisi, Robert F. Service (Pacific NW), Erik Stokstad, Emily Underwood; **web developer** Daniel Berger; **social media strategist** Meghna Sachdev; **intern** Lizzie Wade; **contributing correspondents** John Bohannon, Jon Cohen (San Diego, CA), Ann Gibbons, Sam Kean, Eli Kintisch, Andrew Lawler, Mitch Leslie, Charles C. Mann, Virginia Morell, Gary Taubes; **copy editors** Melissa Raimondi, Kara Estelle; **administrative support** Scherraine Mack; **bureaus** San Diego, CA: 760-942-3252, FAX 760-942-4979; Pacific Northwest: 503-963-1940

Production director Wendy K. Shank; **assistant manager** Rebecca Doshi; **senior specialists** Steve Forrester, Christopher Redwood, Anthony Rosen; **Preflight director** David M. Tompkins; **manager** Marcus Spiegler; **specialists** Jason Hillman, Tara Kelly

Art director Yael Fitzpatrick; **associate art director** Laura Creveling; **senior illustrators** Chris Bickel, Katharine Sutliff; **illustrator** Yana Hammond; **senior art associates** Holly Bishop, Preston Huey; **art associates** Kay Engman, Garvin Grutlón, Chrystal Smith; **photo editor** Leslie Blizard

***Science* International**

Europe (science@science-int.co.uk) **editorial: international managing editor** Andrew M. Sugden; **senior editor/commentary** Julia Fahrenkamp-Uppenbrink; **senior editors** Caroline Ash, Stella M. Hurtley, Ian S. Osborne, Peter Stern; **associate editor** Maria Cruz; **contributing editor** Helen Pickersgill; **editorial support** Rachel Roberts, Alice Whaley; **administrative support** Janet Clements, Jenny Hinson, John Wood; **news: deputy news editor, u.k.** Daniel Clery; **contributing editor, europe** Martin Enserink; **contributing correspondents** Michael Balter (Paris), Kai Kupferschmidt (Berlin), Gretchen Vogel (Berlin)

Asia Japan Office: Asca Corporation, Tomoko Furusawa, Rustic Bldg. 7F, 77 Tenjin-cho, Shinjuku-ku, Tokyo 162-0808, Japan; +81 3 6802 4616, FAX +81 3 6802 4615, inquiry@sciencemag.jp; **contributing editor, asia** Mara Hvistendahl [China: mhvisten@aaas.org]; **contributing correspondents** Dennis Normile [Japan: +81 (0) 3 3391 0630, FAX +81 (0) 3 5936 3531; dnormile@gol.com]; Hao Xin [China: cindyhao@gmail.com]; Pallava Bagla [South Asia: +91 (0) 11 2271 2896; pbagla@vsnl.com]

EXECUTIVE PUBLISHER **Alan I. Leshner**
PUBLISHER **Beth Rosner**

Fulfillment Systems and Operations (membership@aaas.org); **customer service supervisor** Pat Butler; **specialists** LaToya Casteel, Michelle Ofordire, April Marshall; **manager, data entry** Mickie Napoleoni; **data entry specialists** JJ Regan, Jaimee Wise, Fiona Gildin

Business Operations and Administration director Deborah Rivera-Wienhold; **business systems and financial analysis director** Randy Yi; **manager, fulfillment systems** Frits Buningh; **systems analyst** Nicole Mehmedovich; **manager, business analysis** Eric Knott; **manager, business operations** Jessica Tierney; **business analysts** Cory Lipman, Celeste Troxler; Christine Wehrli; **financial analyst** Jeremy Clay; **rights and permissions: administrator** Emilie David; **associate** Elizabeth Sandler; **marketing director** Ian King; **marketing managers** Alison Chandler, Julianne Wielga, Justin Sawyers; **marketing associates** Mary Ellen Crowley, Elizabeth Sattler, Rebecca Riffkin; **senior marketing executive** Jennifer Reeves; **director, site licensing** Tom Ryan; **director, corporate relations** Eileen Bernadette Moran; **senior publisher relations specialist** Kiki Forsythe; **publisher relations manager** Catherine Holland; **publisher relations, eastern region** Keith Layson; **publisher relations, western region** Ryan Rexroth; **customer relations manager** Iquo Edim; **marketing manager** Christina Schlecht; **marketing associates** Paulina Curto, Mitchell Edmund; **customer relations analysts** Simon Chong, Lana Guz; **electronic media: director** Lizabeth Harman; **assistant manager** Lisa Stanford; **production specialists** Antoinette Hodal, Nichele Johnston, Lori Murphy, Kimberly Oster; **web and new media: senior project manager** Trista Smith, **project leader** Luke Johnson **computer specialists** Walter Jones, Kai Zhang, **web developer** Chris Coleman; **program director, aaas member central** Peggy Mihelich

director, global collaboration, custom publications, advertising Bill Moran
editor, custom publishing Sean Sanders: 202-326-6430
assistant editor, custom publishing Tianna Hicklin 202-326-6463
associate director, collaboration, custom publications/china/taiwan/korea/singapore Ruolei Wu +86-1367-101-5294

Product (science_advertising@aaas.org); **midwest** Rick Bongiovanni: 330-405-7080, FAX 330-405-7081; **east coast/e. canada** Laurie Faraday: 508-747-9395, FAX 617-507-8189; **west coast/w. canada** Lynne Stickrod: 415-931-9782, FAX 415-520-6940; **uk europe/asia** Roger Goncalves: TEL/FAX +41 43 243 1358; **japan,** Makiko Hara: +81 (0) 3 6802 4616, FAX +81 (0) 3 6802 4615; ads@sciencemag.jp; **china/taiwan** Ruolei Wu: +86 1367 1015 294 rwu@aaas.org

worldwide associate director of *Science* Careers Tracy Holmes: +44 (0) 1223 326525, FAX +44 (0) 1223 326532

Classified (advertise@sciencecareers.org); **u.s.: east coast/west coast/south central/south america** Tina Burks: 202-326-6577; **midwest/canada/industry** Allyson Rosen: 202-326-6578; **sales administrator** Marci Gallun; **europe/row sales** Lucy Nelson; **sales assistant** Kelly Grace; **japan** Yuri Kobayashi +81 (0)90-9110-1719; careerads@sciencemag.jp; **china/taiwan** Ruolei Wu: +86 1367 1015 294 rwu@aaas.org; **advertising support manager** Karen Foote: 202-326-6740; **advertising production operations manager** Deborah Tompkins; **senior production specialist/graphic designer** Amy Hardcastle; **production specialist** Yuse Lajiminmuhip; **senior traffic associate** Christine Hall; **sales coordinator** Shirley Young; **marketing manager** Allison Pritchard; **marketing associate** Aimee Aponte

AAAS Board of Directors retiring president, chair Nina V. Fedoroff; **president** William H. Press; **president-elect** Phillip A. Sharp; **treasurer** David E. Shaw; **chief executive officer** Alan I. Leshner; **board** May R. Berenbaum, Bonnie L. Bassler, Stephen L. Mayo, Raymond Orbach, Julia M. Phillips, Sue V. Rosser, David D. Sabatini, Inder M. Verma

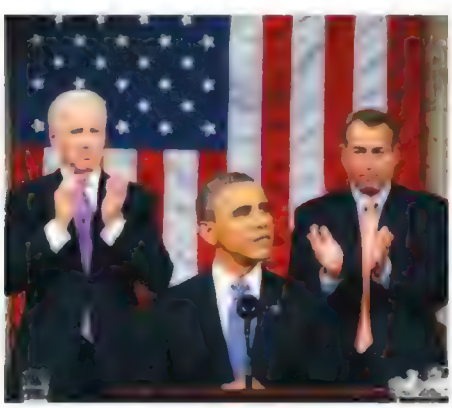
AAAS
ADVANCING SCIENCE, SERVING SOCIETY

AROUND THE WORLD

Washington, D.C. **1**

Science Gets Nod In State of the Union

Many research advocates were smiling after President Barack Obama's annual State of the Union address on 12 February. Obama called on Congress to take steps to thwart climate change, but said he'd act alone if lawmakers didn't. He also asked Congress to protect biomedical research funding from budget



cuts, create 15 "hubs" to pursue innovative manufacturing technologies, and expand "investments in science and innovation" to levels "not seen since the height of the space race." He offered few details, however, on how to fund the initiatives, many of which are unlikely to get through the Republican-controlled House of Representatives. **http://scim.ag/SOTUscience**

Science LIVE

Join us on Thursday, 28 February, at 3 p.m. EST for a live chat on the **science of dog cognition**. What are we learning about the canine mind? **http://scim.ag/science-live**

Washington, D.C. **2**

Flu Funding Rules Finalized By Government

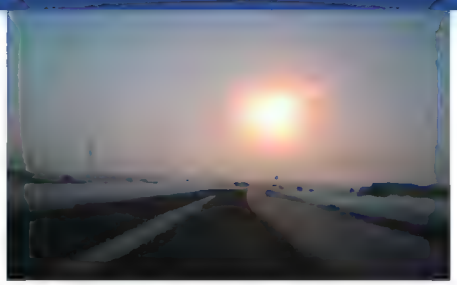
Researchers interested in creating potentially dangerous variants of the H5N1 avian influenza virus now have some new hurdles to jump if they want to get funding from the U.S. government. Officials from the Department of Health and Human Services (HHS) and its National Institutes of Health this week announced in *Science* that they have finalized a framework for reviewing proposals to alter the bird virus in ways that would allow it to move between mammals. The new funding rules were incited by a global controversy that erupted in late 2011 after two teams created such transmissible H5N1 viruses, raising fears of a deadly human pandemic (*Science*, 4 January, p. 16). The framework sets out seven criteria, including one that requires showing that proposed engineered viruses could also "be created through a natural evolutionary process." Some researchers argue such forecasting is impossible. Just a handful of studies are expected to fall under the new rules, which HHS officials say they will regularly review and revise as needed. **http://scim.ag/flufram**

Western Siberia **3**

Meteor Puts on Siberian Show

Last week, a blast from the heavens that shattered windows across Western Siberia, injuring more than a thousand people, drove home planetary scientists' words of warning about the incessant rain of cosmic debris. Before it exploded in the air on 15 February and its meager remains hit the ground, our planet's latest celestial visitor was perhaps 15 meters across.

This latest asteroid was far smaller than



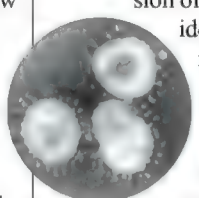
the infamous dinosaur killer, notes planetary scientist Paul Chodas of NASA's Jet Propulsion Laboratory in Pasadena, California. (Scientists estimate that was about 10 kilometers in diameter.) The newcomer, however, may have been the biggest observed since the 1908 Tunguska object, estimated at about 40 meters in size, exploded in the air and leveled 2000 kilometers of Siberian forest.

A dinosaur-killer-size asteroid comes every few hundred million years on average and a Tunguska-size one about every 1200 years, but it's hard to judge the frequency of objects like the latest one, Chodas says. Once a century on average might not be far from reality. **http://scim.ag/Siberiamet**

Birmingham and Manchester, U.K. **4**

People May Spread Novel Virus

A patient sick with a novel coronavirus infected two family members, marking the clearest signs of person-to-person transmission of the new virus, originally identified last year in a man in Saudi Arabia. NCoV, as it is known, is related to the SARS virus that caused a deadly 2003 outbreak, so health officials have been tracking it closely to see how dangerous it might be.



The first nine cases were diagnosed last year in people who lived in or had traveled to the Middle East; five patients died. Because there was no clear evidence that their con-

THEY SAID IT

"The high seas are owned by everyone but their governance and management are inadequate."

—José María Figueres, former president of Costa Rica and a founding member of the Global Ocean Commission, a new independent body aiming to advise the United Nations on how to protect biodiversity in ocean waters outside national jurisdictions. **http://scim.ag/high_seas**

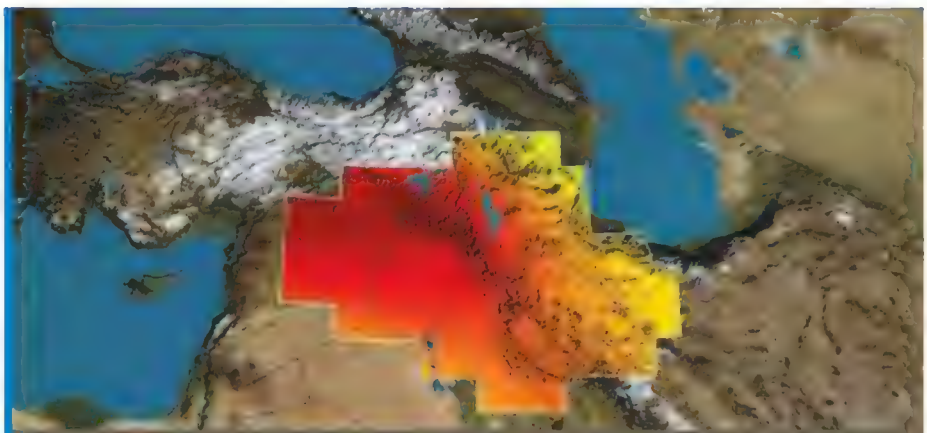
CREDITS (TOP TO BOTTOM): NASHA GAZETA, WWW.NG.KZ/AP PHOTO; FRED MURPHY/CDC/WIKIPEDIA; REX FEATURES VIA AP IMAGES

NOTED

>The U.S. Food and Drug Administration has followed European regulators by approving an artificial retinal implant for the first time. Designed to help people blinded by a genetic condition that destroys light-sensitive cells in the retina, the ARGUS II Retinal Prosthesis System, made by the firm Second Sight, incorporates camera-equipped goggles that transmit low-resolution visual information through a multielectrode array implanted into a person's eye.

tacts became infected, experts thought that each patient could have picked up the virus from an animal host.

On 15 February, the U.K. Health Protection Agency (HPA) announced that three U.K. residents in the same family, only one of whom had recently traveled to the Middle East and Pakistan, were infected. One died on 17 February. Still, HPA emphasized that there is no cause for alarm, noting that if the virus transmitted easily, many more cases would have been seen by now.



Drying Out the Cradle of Civilization

It's long been on the dry side along the Tigris and Euphrates rivers, but new satellite data paint a picture of humans draining the region's meager water resources at an alarming rate. By measuring subtle changes in the pull of gravity over parts of Turkey, Syria, Iraq, and Iran from 2003 through 2009, NASA's twin GRACE satellites have revealed a dramatic loss of about 90 cubic kilometers of ground water (reds are largest losses), as reported last week in *Water Resources Research*. Farmers and other water users struck by a 2007 drought apparently had to withdraw water from wells faster than rain could replenish it.

Geneva, Switzerland **5**

Global Health Needs Better Numbers

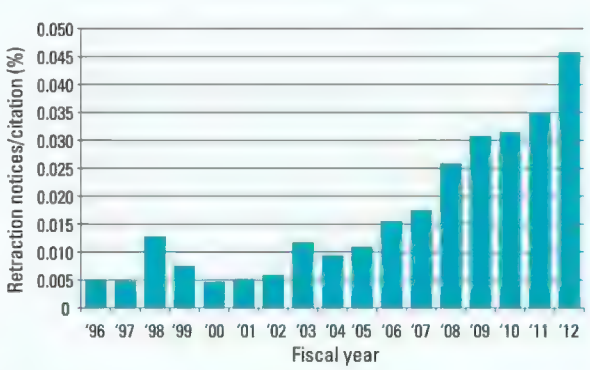
Stay tuned for GBD 2.0, a follow-up to the massive Global Burden of Disease (GBD) 2010 study released in December. That was the largest-ever effort to estimate how many people across the globe suffer and die from nearly 300 diseases and injuries. At a meeting at the World Heath Organization (WHO) last week, GBD scientists said that when they release their country-level estimates next month, they'll also unveil plans to develop a new version of the study, to be updated yearly.

The meeting brought together 60 leaders in the field of global health statistics to discuss how to improve the world's ability to measure the sick and the dead. Knowing how many people suffer from which maladies is crucial for effective public health policies, but in many regions data are sparse. "Where disease burden is greatest, our capacity to measure trends doesn't exist," said Margaret Chan, director-general of WHO.

Scientists use computer modeling to fill in data gaps, but the models are so complex that it's difficult for outsiders to check the estimates. Recent numbers, including some in the GBD study, have sparked controversy (*Science*, 14 December 2012, p. 1414). The meeting diffused some of the tensions, and participants agreed to more sharing of data and methods, as well as calling for investment in better data-collecting.

http://scim.ag/healthstats

Biomed Retractions Climb



Biomedical retraction notices adjusted for the number of papers published annually continue to rise, according to a new analysis in the newsletter of the U.S. Office of Research Integrity (ORI). The increase was especially sharp recently, with 348 retraction notices in 2012 versus 254 in 2011. ORI investigator John Krueger, who did the analysis, theorizes that one reason for the increase may be new technologies that help scientists more easily detect inconsistencies, particularly in published images.

BY THE NUMBERS

$134 billion Estimated reduction in health care costs for California residents over a span of nearly 20 years, linked to its state tobacco control program, according to a new study in *PLOS ONE*.

$140 Claimed return on investment, noted by President Barack Obama in his State of the Union speech, for every U.S. dollar spent on the Human Genome Project.

CREDITS (TOP TO BOTTOM): NASA/UNIVERSITY OF CALIFORNIA, IRVINE/NCAR; (SOURCE) ORI NEWSLETTER/HHS AND JOHN KRUEGER

AAAS | 2013 ANNUAL MEETING

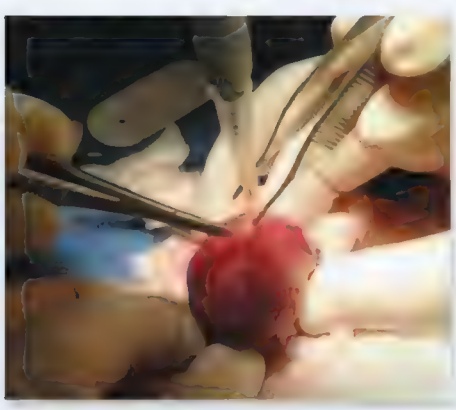
The AAAS annual meeting, held in Boston from 14 to 18 February, attracted a total of almost 10,000 participants, including more than 900 journalists and 3647 people who participated in Family Science Day activities. The following are snapshots from the meeting. For more coverage, including reports from sessions, podcasts, Q&As, video chats, and a new feature called "6-Second Science," go to **http://scim.ag/aaas_2013**.

FINDINGS

Mussel-Inspired Glue Seals Membranes

When it comes to hanging on tight, the lowly mussel has few rivals in nature. Now, researchers have used the mollusk's tricks to develop a biocompatible glue that may one day be used in a variety of medical applications.

Materials scientist Phillip Messersmith of Northwestern University in Evanston, Illinois, reported that he and his colleagues have



created a synthetic, threadlike polymer and attached a synthetic form of an amino acid called DOPA that's abundant in the mussel's glue, to the thread's tips. This DOPA-decorated thread could "more or less recapitulate the central properties of mussel adhesion," Messersmith said. In recent, unpublished experiments, Martin Ehrbar's group at University Hospital Zurich, collaborating with Messersmith's team, used the adhesive to seal a 3.5-mm hole in the fetal membrane of rabbits (pictured). Without the glue, only 40% of the fetal rabbits survived the surgery, but with the glue, 60% did. Messersmith and his colleagues hope that mussel-inspired adhesives will eventually be used to seal up arteries, and help fetal surgeons repair a serious birth defect known spinal bifida caused by an opening in the tissue surrounding the spinal cord.

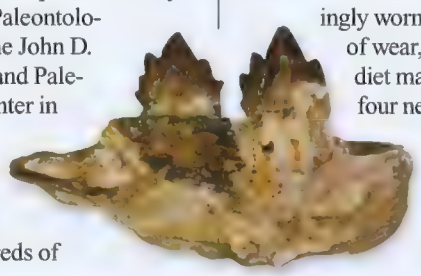
Looking for Life on Mars? Dig Deeper

A martian meteorite that arrived on Earth 12,000 years ago is strengthening the idea that life on Mars—if it exists—is buried deep. The meteorite, found in 1979, contains nitrate and perchlorate, which was a source of energy for some microbes in the early Earth's oxygen-free atmosphere. But did those chemicals originate on Mars or creep in during the meteorite's 12,000-year tenure on Earth? Now that question has been solved: They're definitely from Mars, scientists reported at the meeting.

What's more, new analyses of data from the Curiosity rover's predecessor, Phoenix, suggest that the Red Planet's surface also contains oxychlorines, which react speedily with organic matter if water is present. So if scientists find any organics on Mars, that means the planet is very, very dry—at least nowadays—which isn't good news for life. But life has been found in all sorts of deep, dark, inhospitable spots on Earth, said lead author and planetary scientist Samuel Kounaves of Tufts University in Medford, Massachusetts. "We had to go pretty far down. So if Earth is any guide, we've got a long way to go."

New Whale Species Unearthed in California

Chalk yet another fossil find up to roadcut science. Thanks to a highway-widening project in California's Laguna Canyon, scientists have identified several new species of early toothed baleen whales. Paleontologist Meredith Rivin of the John D. Cooper Archaeological and Paleontological Curation Center in Fullerton, California, who presented the finds at the meeting, said the Laguna Canyon outcrop contained hundreds of marine mammals that lived 17 million to 19 million years ago. Among the finds, she said, were four newly identified but still unnamed species of toothed baleen whale—a type of whale that scientists thought had gone extinct 5 million years earlier.



The four new species are the youngest toothed whales yet discovered. Three of the fossils, including these teeth, belong to the genus *Morawanocetus*, which hadn't been seen before in California. The fourth—dubbed "Willy"—was bigger than the three *Morawanocetus* fossils. Its teeth were surprisingly worn—and based on the pattern of wear, Rivin said Willy's favorite diet may have been sharks. These four new species aren't ancestral to any of the living whales, but they could represent transitional steps on the way to the toothless mysticetes.

The Swift Planet, in Vivid Color

Constructed of thousands of images taken by NASA's MESSENGER spacecraft, this false-color image of Mercury, released this week at the meeting, is the first global color map of the planet. It's a vast improvement over the last global map of Mercury, composed of images from NASA's Mariner 10 spacecraft, which flew by the planet in the 1970s but mapped only about 50% of its surface. The colorful map highlights compositional differences in the surface rock: light-blue fresh impact craters, tan plains of lava, and medium- and dark-blue areas of a low-reflectance material (thought to be a dark, opaque mineral).

CREDITS (TOP TO BOTTOM): NASA/JOHNS HOPKINS UNIVERSITY APPLIED PHYSICS LABORATORY/CARNEGIE INSTITUTION OF WASHINGTON; MARTIN EHRBAR/UNIVERSITY HOSPITAL ZURICH; JOHN D. COOPER ARCHAEOLOGICAL AND PALEONTOLOGICAL CENTER

Economics Dims Coal's Power

Some critics of President Barack Obama and the Environmental Protection Agency claim that a federal "war on coal" has led to shuttering U.S. coal-fired power plants and putting coal miners out of work. Not so, according to a new analysis of the coal industry by David Schlissel of the Institute for Energy Economics and Financial Analysis in Belmont, Massachusetts. Instead, he said, coal is losing its battle with other power sources, mainly natural gas, due to simple economics.

Over the past 6 years the share of electricity generated by coal plants has fallen from 50% to 38%. Plans for more than 150 new coal-fired power plants have also been canceled since the mid-2000s. Schlissel, who has served as a paid expert witness at state public utility board hearings for both utilities and advocacy groups that oppose coal plants, found that sharply rising costs in the construction business have made new plants very expensive, and many existing plants face growing maintenance costs—60% of the nation's coal plants are more than 40 years old. Meanwhile, the price of natural gas has plummeted in recent years. "I don't think there's any question" that coal is losing on its economic merits, says Melissa Ahern, an economist at Washington State University, Spokane. In addition to the factors Schlissel cited, the costs of shipping coal by train and truck are large and rising, which adds significantly to the fuel's cost, she notes.

Ancient Pee Provides Clues To Africa's Past

When it comes to peering into Africa's climate past, the ancient homes of hyraxes are number one. Paleoclimatologists typically dig up muddy core samples and analyze their pollen content for clues to long-ago weather, but parts of southern and central Africa are too dry to preserve such evidence. Enter the rock hyrax, *Procavia capensis* (inset), a furry mammal that looks like a large groundhog but is actually a distant cousin of the elephant. Hyrax colonies use the same rock shelters for generation after generation, depositing pollen, calcium remnants, charcoal particles, stable isotopes, and other detritus in their urine. "You can turn a 2-meter pile of pee into a very nice section which you can bring back to the lab," Brian Chase, a geographical scientist at the University of Montpellier in France (pictured), told the audience. "These are very high-resolution [climate] records." Most climate models predict arid conditions in southern Africa 12,000 years ago, but the pollen content of hyrax urine from that period indicates that they ate grasses, which flourish in wetter conditions, he reported.

CREDITS (TOP TO BOTTOM): UNIVERSITY OF LEICESTER; (INSET) ARIKK/WIKIMEDIA COMMONS; WIKIMEDIA COMMONS; (RIGHT) *SCIENCE* (2)

NEWSMAKERS

Three Q's

Tegmark

For physicist **Max Tegmark** of the Massachusetts Institute of Technology in Cambridge, it's not enough to say that math governs our universe; he thinks reality itself is a mathematical structure. Tegmark described his ideas in a symposium called, "Is Beauty Truth?"

Q: What makes a mathematical theory beautiful?

M.T.: The beautiful mathematical regularities that have been uncovered have typically been unifications, where instead of having one mathematical description for this and a different one for that, we realize there's a single mathematical structure that encompasses all of it. So for me, it would be a natural conclusion if there's a single mathematical structure that is our reality, and all of the mathematical structures that we've discovered before are part of this more beautiful whole.

Q: What do you mean: the universe is a mathematical structure?

M.T.: Right now, I'm eating an orange. Why do [its] molecules have their properties? Because they're made of atoms put together in a certain way. Why do the atoms have those properties? Because they're made of quarks and electrons. [But], all the properties electrons have are purely mathematical. It's just a list of numbers. In fact, there's no evidence that there's anything at all in our universe that is not mathematical.

Q: You call this idea a fundamentally optimistic way of looking at reality. Why?

M.T.: If the mathematical universe hypothesis is true, we can actually learn things about the parts of our universe we can't see or visit. Not with a telescope but with a pencil—and a lot of ingenuity. If the hypothesis is false, we're ultimately going to hit a roadblock beyond which we just cannot proceed. Whereas if I'm right, there is no roadblock. The road ahead is open, and our future understanding is really only limited by our imagination.

SIX-SECOND SCIENCE

Science challenged meeting attendees to describe their research in 6 seconds, and we captured the results with an iPhone video app called Vine. Here's a selection; you can view all the videos at **http://scim.ag/6SecSci.**

"Language co-evolved in the human lineage with technological know-how and social cooperation."

—Harvard University psychologist Steven Pinker

"Here's a star, here's a planet, the magnetic fields link up, the magnetic energy creates star spots on the star."

—Boston University astronomer Mark "Astro" Zastrow

"I'm proposing a tool to communicate uncertainty in river-stage forecasts to emergency managers."

—Civil engineer Frauke Hoss of Carnegie Mellon University in Pittsburgh, Pennsylvania

"I'm Jeremy deSilva from Boston University, and 26 bones in the human foot is not an intelligent design."

eppendorf & Science

PRIZE FOR NEURO BIOLOGY

2012 Winner
Dr. Marlene R. Cohen
Assistant Professor
University of Pittsburgh

Application Deadline
June 15, 2013

Call for Entries

Eppendorf & Science Prize for Neurobiology
The annual Eppendorf & Science Prize for Neurobiology, an international award, honors young scientists for their outstanding contributions to neurobiological research. The winner and finalists are selected by a committee of independent scientists, chaired by Science's Senior Editor, Dr. Peter Stern. To be eligible, you must be 35 years of age or younger.

You could be next to win this prize and to receive
> Prize money of US$25,000
> Publication of your work in Science
> Full support to attend the Prize Ceremony held in conjunction with the Annual Meeting of the Society for Neuroscience in the USA
> An invitation to visit Eppendorf in Hamburg, Germany

It's easy to apply!

Learn more at: www.eppendorf.com/prize

eppendorf

Science AAAS

Primed for action. Palmer Station in Antarctica hosts one of CTBTO's 32 radionuclide detectors (*inset*).

NUCLEAR WEAPONS

North Korea's Blast Poses Riddles and Challenges

In a closed-door meeting in Beijing in the summer of 2011, Chinese and South Korean experts huddled over photos of centrifuges at North Korea's Yongbyon nuclear complex. The machines' existence had come to light the previous November, when Stanford University metallurgist Siegfried Hecker visited the facility; the Yongbyon scientists told Hecker that the cascade was for enriching uranium for a future power reactor. The revelation accentuated fears—long denied by North Korea—that it was engaged in a clandestine effort to produce highly enriched uranium (HEU) for bombs. Chinese experts had withheld judgment—until they saw the photos from South Korea, says a Chinese physicist present at the meeting who requested anonymity due to the issue's sensitivity in China. "My colleagues said, 'Wow, these machines are the real thing, not toys.'"

The same might be said of North Korea's self-proclaimed "nuclear deterrent." Last week, the nation conducted its third nuclear weapons test, detonating a bomb that packed a far more powerful punch than its previous efforts. Speculation is swirling around whether Korea tested a plutonium device, as analysts presume it has done twice previously, or a uranium bomb. A uranium weapon "would be a game-changer," says Mark Hibbs, an analyst based in Berlin for the Carnegie Endowment for International Peace, because North Korea has ample deposits of the element and could build up an HEU stockpile for bombs or for sale. As *Science* went to press, hopes were fading that nuclear sleuths would sample short-lived radioactive fission products venting from the underground test site and reveal the nature of the bomb.

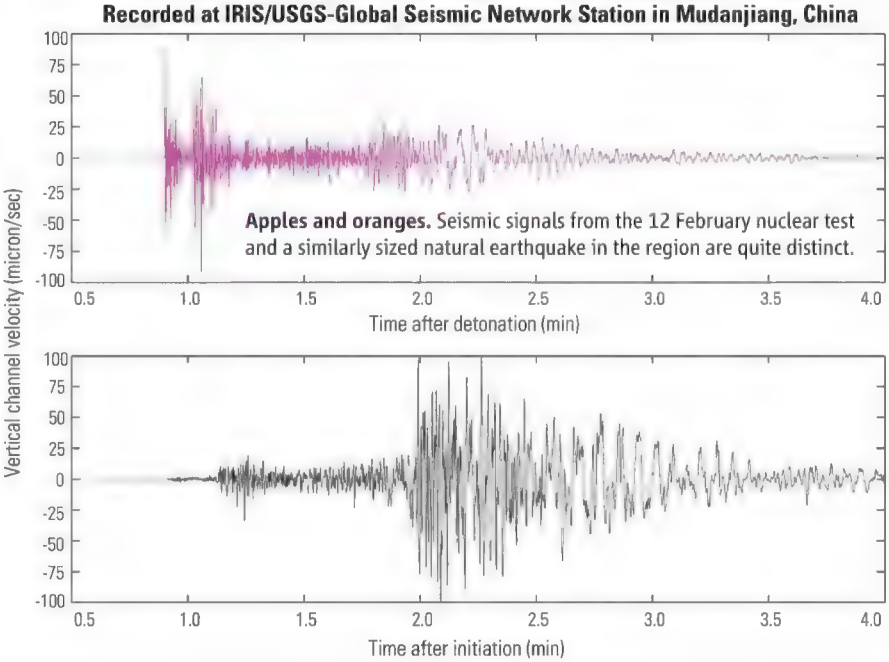
Experts do not doubt that the latest test was the real McCoy, and that the nuclear renegade—North Korea dropped out of the Treaty on the Non-proliferation of Nuclear Weapons in 2003—is making strides toward a potent arsenal. The threat is limited until Korea acquires a sophisticated means to deliver a nuclear warhead to a target. But that may only be a matter of time. Last week, North Korea claimed that it had tested "a smaller and light A-bomb unlike the previous ones," implying progress toward miniaturizing a warhead that could be put on a missile.

Heralding North Korea's latest nuclear handiwork was a magnitude-5.1 earthquake just before noon local time on 12 February in the northeastern corner of the country. The shallow tremor was centered in rugged hills where the first two nuclear tests were carried out in 2006 and 2009. On all three occasions, the seismic waves were consistent with an explosion: a burst of P-wave energy from compression and less shearing, or S-wave, energy than would be seen in a natural earthquake (see graphic, below).

North Korea's 2006 test had an estimated yield of around 1 kiloton: a fizzle compared to bombs in modern arsenals. That might have raised suspicions that Korea had exploded a chemical bomb, rather than a nuclear charge, had sensors not detected the smoking gun of an underground test: xenon radioisotopes seeping from the site. Although the Comprehensive Nuclear-Test Ban Treaty (CTBT) has not come into force, a monitoring network is already in place to verify compliance. The network includes 32 stations for measuring gamma rays from four radioxenon isotopes—Xe-133, Xe-133m, Xe-131m, and Xe-135. The three stations closest to the test site are in Beijing; Takasaki, Japan; and Ussuriysk, Russia.

The second test in 2009 released six times more energy than the first. Still, it was at best one-half as powerful as the bombs dropped on Japan at the end of World War II. "No other nuclear armed state spent so much time struggling with yield," says Carnegie

2013 M5.1 Nuclear Test (Top) and M5.0 Regional Earthquake (Bottom) Recorded at IRIS/USGS-Global Seismic Network Station in Mudanjiang, China



Apples and oranges. Seismic signals from the 12 February nuclear test and a similarly sized natural earthquake in the region are quite distinct.

CREDITS (TOP TO BOTTOM): CTBTO (2); ANDY FRASSETTO/INCORPORATED RESEARCH INSTITUTIONS FOR SEISMOLOGY

arms control expert Li Bin. "That's strange and indicates problems with their weapon design." He notes that other nations reached 10 to 20 kilotons on their first test. Estimates for last week's explosion are all over the map, ranging from 6 to 40 kt.

Although it's implausible that the 2009 test was anything but nuclear, sensors failed to register radioxenon wafting from the test site. One hypothesis is that the explosion's heat melted enough rock surrounding the cavity to trap radioactive gases. Another possibility is that "the North Koreans sealed the explosion chamber better the second time," says nuclear physicist Anders Ringbom, deputy research director at the Swedish Defence Research Agency in Stockholm.

Online
sciencemag.org
Podcast interview with author Richard Stone (http://scim.ag/pod_6122).

North Korea may have gotten lucky again. CTBT monitors have come up empty so far, and no reports have leaked of positive hits from a U.S. Air Force plane dispatched from Japan to sample for noble gases. Time is running out: Half-lives of the four radioxenon isotopes of interest range from 9 hours to 12 days. Measuring radioxenon isotopic ratios as soon as possible after the test is critical for identifying the fissile material: "There are different telltale signatures for HEU and plutonium devices," Hecker says.

The type of material used in last week's test has huge implications. North Korea has enough plutonium for only four to eight bombs, Hecker estimates. The plutonium was extracted from spent uranium fuel rods from a Yongbyon reactor whose cooling tower was demolished in 2008. It would take years to restart the reactor and reprocess more spent fuel, meaning that North Korea has a finite plutonium supply for the foreseeable future. But it could ramp up HEU production, unseen, by dispersing centrifuges to easily hidden warehouses.

As scientists grapple with the uncertainties, diplomats are struggling over a response. All eyes are on China, which expressed pique over the test but has resisted measures that could undermine the North Korean regime. This time around, China may feel compelled to take a tougher line to maintain the balance of power in Northeast Asia. It's in a tense standoff with Japan over disputed islands; Japan's conservative government could use the Korean test as another reason to up its defense spending. And South Korea and the United States are negotiating a nuclear cooperation agreement that will succeed an accord due to expire in March 2014. The United States has so far resisted the South's demands to be allowed, like Japan, to reprocess spent fuel from its power reactors. North Korea's latest blast—and China's response—could test U.S. resolve in deterring South Korea from separating out plutonium. Adding everything up, Hibbs says, "we're moving in a very dangerous direction."

–RICHARD STONE

PHYSICS

Congress Tries Again to Head Off Looming Helium Crisis

For the second time in 8 months, a bill has been introduced in Congress that would prevent an acute but wholly self-inflicted shortage of helium from striking later this year. The shortage would hamstring research in a variety of fields. "Unless Congress takes swift action, the U.S. will float off the helium cliff," said Representative Doc Hastings (R–WA), chair of the House of Representatives Natural Resources Committee and co-sponsor of the bipartisan bill, during a hearing this week.

Here's the problem. Since 1996, the U.S. government has been selling off its vast reserve of helium, a legacy of the Cold War, which is held in an underground reservoir near Amarillo, Texas. Those sales supply 42% of the crude helium consumed in the United States and 35% of the crude helium consumed globally. But by law, they will continue only until the Bureau of Land Management (BLM), which controls the reserve, recoups the $1.4 billion cost of developing the reserve. BLM will reach that break-even point by the end of this fiscal year, 30 September. At that point, BLM will have no authority to sell the remaining helium unless Congress acts.

Helium is a byproduct of natural gas production, although only some gas wells produce helium at high enough concentrations to make it worth capturing. Much helium is lost at the wellhead. So even with the BLM reserve pumping, helium is in short supply, says Samuel Aronson, former director of Brookhaven National Laboratory in Upton, New York, and vice president of the American Physical Society. "A number of the national labs are now receiving about two-thirds of their allocations of helium," he testified.

Helium has myriad uses. The only element that remains a liquid at absolute zero, it cools the superconducting magnets in MRI machines and particle accelerators such as Europe's Large Hadron Collider. Many condensed matter physicists rely on liquid helium as a coolant, and helium balloons take experiments aloft. The gas is essential for producing microchips and optical fibers.

Ubiquitous. Laboratories around the world depend on helium refrigerators like this one (blue).

CREDIT: NEAL E. STALEY/MASSACHUSETTS INSTITUTE OF TECHNOLOGY

Handwringing over the helium supply is nothing new. In 1996, Congress instructed BLM to sell off the reserve at a steady rate by 2015 and at a price set to recover the reserve's debt. BLM's price started out above the market price for refined helium but by 2010 had sunk below it, and a report by the National Academies found that BLM's sales were suppressing the price of helium. The report recommended BLM charge a market price, draw the helium out in a way that would maximize recovery, and protect supply for federal users including researchers.

The new bill addresses those concerns. It would continue current sales for another year and then, in a second phase lasting until about 2020, sell the helium through biannual auctions. Finally, when 85 billion cubic liters of gas were left—less than a tenth of the original reserve—the bill would allow sales only to federal users, including research grant holders. A largely identical Senate bill that would have set prices by polling vendors' rates for helium instead of by auction died in committee last year.

Even if the bill becomes law, the helium reserve will be gone within roughly 15 years. Rush Holt (D–NJ), a physicist by training, wondered whether the government should keep a reserve for unforeseen uses. "I'm really concerned that some critical [future] uses will face short supply because we didn't think of them," he said. No other legislator expressed support for that idea.

–ADRIAN CHO

NOAA

Warning Issued for Looming Data Gap From Fleet of Weather Satellites

U.S. policymakers are becoming increasingly alarmed about a looming gap in critical weather forecasting data collected by the nation's satellites.

Last week, the Government Accountability Office (GAO), Congress's watchdog arm, added the data gap to a short list of "high risk" issues that deserve special attention. In anticipation, some lawmakers last month quietly set aside $111 million to help close the gap as part of a $50 billion aid package for areas hit hard last fall by Superstorm Sandy. But some scientists and others are worried about how those funds will be used.

The flurry of activity highlights a growing sense "that there is going to be a gap, and it could be serious," says analyst David Powner of GAO. The report estimates that the break in the data could extend for 17 to 53 months, and begin as soon as late 2014 under a worst-case scenario. "We've been warning about it for years," he says, "and now it's happening."

Blind spot. Weather forecasters are facing a 17-month gap in data collected by U.S. polar-orbiting satellites. The gap could be as long as 53 months if the Soumi satellite fails early (*left*) or the JPSS-1 orbiter (*right*) is delayed.

The gap is the product of a messy, nearly 20-year-old effort to improve the nation's fleet of polar-orbiting satellites, which provide the majority of the data used to develop weather forecasts extending out 3 to 10 days. Bristling with instruments that measure things like cloud cover and atmospheric moisture levels, these spacecraft circle from pole to pole some 14 times a day at a distance of 870 km above Earth's surface. The orbit enables sensors to "see" Earth's entire surface twice daily. Forecasters draw measurements from four primary polar orbiters—three American and one European—that cross the equator at different times of day, meaning their data is typically no more than 6 hours old.

Historically, two agencies—the Department of Defense and the National Oceanic and Atmospheric Administration (NOAA)—have operated the U.S. orbiters with help from NASA. In 1994, however, White House officials tried to reduce duplication and lower costs by creating a single program that would build and operate new orbiters. That joint arrangement collapsed in 2010, however, after years of mismanagement, technical problems, and cost overruns.

The responsibility has now fallen largely on NOAA. The agency's $12.9 billion program, known as the Joint Polar Satellite System (JPSS), includes three major orbiters. The first—named after pioneering American meteorologist Verner Suomi, who died in 1995—was launched in 2011 and is expected to last until late 2016. But Suomi's replacement, known as JPSS-1, isn't expected to be fully operating until early 2018. That leaves a gap of 17 months, analysts say—or longer if Suomi falters or JPSS-1 is delayed.

Suomi's orbit takes it over the equator in the early afternoons, and the loss of that data will "result in less accurate and timely weather forecasts and warnings" of potentially life-threatening "extreme events" such as hurricanes, GAO warned. Some scientists believe the afternoon data played a critical role in predicting the unusual path of Superstorm Sandy last fall and the severity of the 2010 "Snowmageddon" blizzard that buried the eastern United States.

Data to go missing

- cloud cover
- atmospheric temperature and moisture profiles
- fire detection
- ozone later mapping
- radiant energy
- solar irradiance

There's no simple way to fill the gap, say NOAA officials, who in the last few months have ramped up efforts to develop a strategy. The agency has started a Web site (www.osd.noaa.gov), for example, featuring more than 100 ideas that it has received from the public. The proposals include extending the life of instruments currently in orbit, flying temporary sensors on relatively inexpensive platforms such as balloons, aircraft or fleets of microsatellites, and buying substitute data from foreign governments. NOAA already has some data-sharing agreements with Europe, but officials say that expanding the approach could require new diplomatic and security agreements, and possibly expensive technical changes to ground stations and software.

Although NOAA has yet to announce a plan or a price tag for closing the gap, in late January it received $111 million in the Sandy relief bill to establish a "data mitigation gap reserve fund." Legislators didn't specify how NOAA should use the money, however, and the past management problems have raised "some concern about throwing good money after bad," says Kevin Wheeler, vice president of the Consortium for Ocean Leadership, a nonprofit science group based in Washington, D.C.

On 7 February, the consortium and the University Corporation for Atmospheric Research wrote to NOAA and members of Congress suggesting that the U.S. National Academy of Sciences, "or some other external oversight entity," help the agency develop a viable plan. The two groups also urged support for COSMIC-2, a U.S.-Taiwanese plan to launch a constellation of 12 tiny satellites that would harness GPS signals to collect atmospheric data.

NOAA may be facing other data gaps, GAO warns. Two U.S. military satellites scheduled to replace current polar orbiters in the next few years were built in the 1990s and may not operate as long as expected. Europe is also facing challenges in replacing its spacecraft. Those problems could punch holes in early- and mid-morning data streams.

In 2015, according to GAO, NOAA could also experience a 1-year gap in data provided by two geostationary satellites that sit at fixed points over the United States. The agency typically flies three of these craft simultaneously, two active and a backup. But any launch delay could mean a backup won't be available if one of the active craft conks out, as has happened twice since 2008.

–DAVID MALAKOFF

CREDIT: (SOURCE) U.S. GENERAL ACCOUNTABILITY OFFICE

EVOLUTION

How Sweet It Is: Genes Show How Bacteria Colonized Human Teeth

When humans invented farming 10,000 years ago, they weren't the only ones to get a boost from the new starchy diet. Some microbes that had lurked at low levels in the mouths of hunter-gatherers bloomed on the sugary films coating the teeth of farmers who munched cereal grains. Eventually the cavity-causing *Streptococcus mutans*, for one, took root. It adapted to the sweet life, multiplying like a weed and edging out many other species of bacteria. That leaves the modern mouth a depauperate ecosystem, according to two new genetic studies.

Quite a mouthful. Starchy diets left plenty of tartar in medieval English mouths, including a rind around a tooth (*below*), and a chunk so big it was used as a chewing surface (grayish mass, *above*). The tartar trapped bacteria such as this calcified colony (*right*).

The papers come at the problem from two completely different angles: One team analyzed DNA from tartar on ancient teeth, noting changes in the mix of bacterial strains over time. The other zeroed in on the genome of *S. mutans* taken from the mouths of living people, tracking the genetic changes that allowed the bacterium to thrive in sweet, acidic conditions. Taken together, the papers show how cavity-causing bacteria underwent a dramatic population explosion, perhaps about the time humans began to cultivate plants 10,000 years ago.

"This is the first record of how we've upset the bacterial system in the mouth," over thousands of years, says molecular evolutionist Alan Cooper of the University of Adelaide in Australia, lead author of a report published this week in *Nature Genetics*. "The modern mouth is like a landscape nuked with a herbicide, with invasive weeds jumping into vacant niches."

That "herbicide" was starchy sugars, and the "weeds" were sugar-loving *S. mutans* and other bacteria, which are associated with diseases such as dental cavities, diabetes, and cardiovascular disorders. Fossils of humans show surges in cavities and gum disease during the transition to agriculture and the Industrial Revolution (*Science*, 25 May 2012, p. 973). But the new papers are the first to examine genetic change in the bacteria themselves.

Back in the 1980s, bioarchaeologist Keith Dobney of the University of Aberdeen in the United Kingdom realized that fossil teeth harbor a "chemical soup" of bacteria and organic matter trapped inside the tartar, a scaly calcium carbonate deposit. You can actually see the bacteria under a scanning electron microscope, says Dobney, co-author of the *Nature Genetics* paper. But Dobney and Cooper's early efforts to extract bacterial DNA were stymied because there was no way to tell contaminant DNA in the lab from target DNA.

Then, new DNA sequencing methods provided higher resolution sequences, allowing the researchers to differentiate between ancient bacterial DNA in the plaque and both environmental and laboratory contamination by modern strains. The high resolution also allowed Cooper, Dobney, and colleagues to detect the percentage of different species of bacteria within each tartar deposit. The team extracted and sequenced DNA from the teeth of 34 prehistoric human skeletons from Northern Europe, dated from about 7500 years ago to 400 years ago, including the last hunter-gatherers known in Poland and the first farmers found in Germany.

With these data, they charted the spectrum of bacterial species within the human mouth over the past 7500 years. Hunter-gatherers had fewer species that cause cavities and periodontal disease, and different percentages of all 15 phyla of bacteria found in modern teeth, plus some unclassified bacteria. Early farmers showed a sharp increase in bacteria that cause tooth decay, such as a Veillonellaceae strain, and a dramatic surge in *Porphyromonas gingivalis*, which causes periodontal disease.

The first significant sample of *S. mutans* turned up around 4300 years ago in the Bronze Age in Yorkshire, U.K. Overall bacterial diversity, including higher levels of *S. mutans* and *P. gingivalis*, seemed to stabilize through the Bronze Age and medieval period. Then, sometime in the past 400 years, the diversity of bacteria within each tooth dropped sharply once again. Modern samples and preliminary data from the mid-19th century show an oral environment even more dominated by *S. mutans*. Cooper concludes that the change occurred with the Industrial Revolution, about 1850 in England, when cavities also increased and refined sugars entered the diet. "Sugar and flour caused everything to go berserk," he says.

While Cooper and his colleagues were probing the entire bacterial soup in ancient teeth, another international team focused on *S. mutans* alone, using DNA from modern mouths and working back in time. Omar Cornejo, a postdoc at Stanford University in Palo Alto, California, sequenced the genomes of strains of *S. mutans* from the mouths of 57 people around the world. Working with Stanford population geneticist Carlos Bustamante, Cornejo analyzed the distribution of variation in the *S. mutans* genome among individuals. Then he used population genetics models to calculate how long it would take a growing population to generate this amount of diversity. His calculations have very wide margins of error, but suggest that the *S. mutans*

CREDITS: ALAN COOPER (2); KEITH DOBNEY AND DON BROTHWELL/UNIVERSITY OF YORK, ALAN BOYDE/UNIVERSITY COLLEGE LONDON (INSET)

"population started expanding exponentially … 10,000 years ago," which coincides with the start of agriculture in the Near East, the team reports in a paper available online that will appear in the April issue of *Molecular Biology and Evolution*.

Cornejo and his colleagues also tracked which genes changed over time. Out of 1490 genes in *S. mutans*, they identified 14 under positive selection that are thought to be involved in sugar metabolism or acid tolerance. Among 73 genes present in *S. mutans* but absent in other bacterial species, most "seem to be involved in sugar metabolism and resistance to low pH and oxidative stress," Cornejo writes in an e-mail. That would have helped *S. mutans* adapt to "its new niche, the human mouth," as it became rich in starchy sugars.

Taken together, the two papers make a strong case that the oral bacterial community shifted dramatically with the change in diet associated with agriculture. "I find their date of expansion of about 10,000 years very believable," says population geneticist Francois Balloux of University College London who was not part of either study.

The papers don't agree on all points—Cooper's team doesn't detect *S. mutans* in Europe until about 4300 years ago, which is later than the start of agriculture. But that could be the result of sampling too few teeth, Cooper says. Cooper also thinks the shift in the Industrial Revolution, which doesn't appear in Cornejo's study, might be the most dramatic change. Other researchers add that both teams must address issues such as how toothbrushes, fluoride, and antibiotics are affecting bacterial population dynamics.

But it is early days for this work, and researchers have big plans. Dobney says: "If we can mine pathogens and microbes from dental calculus, there's information relevant to a whole boatload of questions about illnesses that started when we began domesticating animals, and modern conditions like diabetes. We're right at the beginning of this."

–ANN GIBBONS

ZOOLOGY

International Arbiter of Animal Names Faces Financial Woes

A rose by any other name might still smell as sweet, but an animal with two scientific monikers can wreak havoc for researchers trying to study it. Since 1895, the International Commission on Zoological Nomenclature (ICZN) has helped ensure animal names are unique and long-lasting, with a panel of volunteer commissioners who maintain naming rules and resolve conflicts when they arise. But the U.K.-based charitable trust that supports all this is slated to run out of money before the year's end—and that could spell trouble. "If the trust ceases to exist it will be very difficult for the commissioners to do their work," says Michael Dixon, chair of the trust's board and director of the Natural History Museum in London. If ICZN disappeared "it would be something akin to anarchy in animal naming."

Knowing the right name for each species is critical for biological research. The ICZN code, which provides guidance for naming new species and resolving conflicts about older ones, is updated periodically. In 2012, for example, after 4 years of debate, ICZN decided to allow species to be named in electronic publications as well as in printed journals. Species-naming records date back centuries, and the concern was that electronic media might not last that long. The commissioners make 30 to 50 judgments a year, some with far-reaching consequences, and the zoological community pays heed.

"Names have to be universal and stable," explains Jonathan Coddington, associate director for science at the Smithsonian Institution National Museum of Natural History in Washington, D.C. There is constant tension, however, between keeping names the same and adjusting them according to new scientific evidence. Consider the popular lab insect, *Drosophila melanogaster*, which some taxonomists want to assign to a new genus because this fruit fly is but a distant relative to some other *Drosophila*. The commission will likely have a hand in the ultimate fate of this name.

Name change? The future of the commission that might help decide the fate of this fruit fly's name is up in the air.

CREDIT: MR. CHECKER/WIKIMEDIA COMMONS

The nonprofit organization that formed in 1947 to raise funds and administer the ICZN code and the journal—the International Trust for Zoological Nomenclature—has weathered other crises. But net income from its journal is only about $47,000 a year, and the trust's annual expenses now top $155,000. So reserves are about to be exhausted, Dixon says.

A few weeks ago, he sent an e-mail plea to directors of natural history museums around the world for emergency relief. In it, he proposed establishing a committee that would come up with a new financial model for the troubled organization. "This is not unlike GenBank," the database of genome sequences that receives government support, Coddington says. "It's the same distributed goods [situation], that everyone needs and nobody wants to pay for."

Botanists, phycologists, and mycologists have for 100 years followed an all-volunteer model. Proposed changes to the botanical code are discussed en masse every 6 years at the International Botanical Congress, and disputes are resolved by volunteer subcommittees. But Dixon says such a model would be tough for zoologists because they don't have an equivalent meeting that's well attended.

Dixon estimates the trust needs $78,000 or more to make it through the year. No single organization may be able to fund it long-term, but a network of 10 or 20 institutions might be able to kick in enough to sustain it, he says.

If support is not forthcoming and ICZN falters, not much would change at first, scientists say. But "communication barriers would start to evolve" over time as names ceased to be systematically assigned and used, says Richard Pyle, a commissioner based at the Bishop Museum in Honolulu. "And that would be tragic."

–ELIZABETH PENNISI

Reversal of Misfortunes

South Africa has the world's worst HIV/TB syndemic; now it's trying to overhaul its response to both diseases

KWAZULU-NATAL PROVINCE, SOUTH AFRICA—On 4 April 2012, an 11-year-old boy named Sandile left his family in Ulundi and traveled 250 kilometers south to his new home at King George V, a hospital in Durban that specializes in the region's most complicated tuberculosis (TB) cases. Like 69% of the other 224 patients at King George V, Sandile was co-infected with HIV, which in all likelihood he acquired from his mother. Records that came with him indicated that she had died, but they did not specify the cause.

Sandile had received a full 6-month course of treatment for TB in 2009. But HIV-infected people remain especially vulnerable to *Mycobacterium tuberculosis*, and a few years later, he again developed an active case of the lung-damaging disease. The doctors here surmise that he didn't receive his medication each day during his second bout, allowing drug-resistant strains of *M. tuberculosis* to flourish. With drug-resistant TB, his treatment at King George V would require 18 months of costlier, less effective, and more toxic "second-line" pills plus painful daily shots of medication for at least 1 year.

Multidrug-resistant (MDR) TB destroyed Sandile's left lung and badly damaged his right one. In September, Sandile was healthy enough to play outside, but Babu Sunkari, the pediatrician in charge of the children's ward, warned that Sandile had nearly died the month before: He had a severe respiratory attack and kidney failure, probably from a combination of TB drug side effects and an infection he picked up at the hospital, further compounded by his HIV infection.

Sandile urgently needed a lung transplant, Sunkari said, yet, like most South Africans without means—which is almost everyone infected with TB here—there was no transplant in the offing. "He doesn't have a hope in hell of surviving," Sunkari said.

Sunkari said he was seeing more and more of these sad cases. Sandile was one of 32 children on the pediatric ward at King George V last September who had MDR or the even more frightening extensively drug-resistant (XDR) TB. "This used to be the normal TB hospital, with one to two cases of MDR," Sunkari said. "Now we're not taking normal cases in the ward."

Children like Sandile "mirror what is happening in the community," says Nesri Padayatchi, who ran the TB program at King George V for 14 years until leaving to do clinical HIV/TB research at Durban's Centre for the AIDS Programme of Research in South Africa (CAPRISA).

South Africa has the world's worst convergence of HIV and TB. It has 5.7 million HIV-infected people, more than any country; the highest rate of TB per capita; and, after Russia, the second highest reported number of MDR TB cases. The big picture is that South Africa has less than 1% of the global population but more than 25% of people dually infected with these diseases—a particularly deadly combination. This alarming rise in coinfection has been dubbed the HIV/TB syndemic, because, for both biological and social reasons such as poverty, the two diseases have synergistic effects, with each making the other worse.

South Africa's HIV/TB syndemic has its roots in a badly broken health system and years of neglect. For TB, a cure has existed for decades, but many sick people in South Africa remain undiagnosed, and those who receive treatment often fail to take their drugs as prescribed, which in turn gives rise to drug-resistant strains. "The reason we have so much MDR and XDR TB here is because implementation of our TB program has sucked," Padayatchi says.

HIV infection cannot be cured, but drugs can effectively thwart disease for decades

CREDIT: MALCOLM LINTON

Double jeopardy. Babu Sunkari shows the lungs of his young patient Sandile, who was infected with both HIV and *M. tuberculosis*.

and also slow the spread of the virus. But from 1999 to 2008, as the AIDS epidemic exploded in the country, then–South African President Thabo Mbeki and his health minister—who for many years questioned whether HIV even caused the disease—dragged their feet when it came to both treatment and prevention.

Now, with President Jacob Zuma at the helm and a new health minister in place, South Africa has made a dramatic turnaround and is at the forefront of efforts to combat these married diseases. "We're trying to integrate HIV/AIDS and TB, to regard them as two sides of the same coin," says Minister of Health Aaron Motsoaledi, who Zuma appointed when he took office in 2009. And to that end, the government has allocated more of its own money and attracted substantial international support to find everyone who needs drugs for HIV and TB, better coordinate their treatment, improve prevention efforts, and ramp up research to figure out how the diseases interact.

But South Africa still has a long way to go, and progress often occurs more slowly than anyone would like.

HIV and TB's disastrous marriage

In the mid-1990s, South Africa followed the lead of many other countries and introduced what's known as DOTS—directly observed therapy, short course—to combat tuberculosis. DOTS used new drugs that led to cures more quickly, and as part of the program, observers watched people swallow their pills each day to make sure they took all of their doses. But just as TB rates began to decline, HIV infections skyrocketed, and by the turn of the century, the two diseases were climbing in parallel (see graph, p. 900).

Salim Abdool Karim, a clinical epidemiologist who heads CAPRISA, helped launch the DOTS program in Hlabisa, a rural municipality not far from Ulundi. At the time, he says, TB mainly afflicted elderly men. "In the mid '90s, a couple of things hit us in a way that we just couldn't grasp—overnight it caught up with us," Karim says. Suddenly, the TB patients were predominantly young and female, a reflection of the fact that the AIDS virus was racing through that population. "All our previous efforts to improve TB control were overwhelmed," he says.

Pre-HIV/AIDS, South Africa's DOTS program—which certainly had operational challenges—made the headway that it did because intact immune systems help contain both illness and spread. *M. tuberculosis* infects one-third of the world's population, but it remains latent in most, typically causing disease only when HIV or other factors compromise the immune system; 90% of people who have latent TB infections never develop active disease in their lifetime. But HIV drastically alters that equation. A person with a latent TB infection who then becomes infected with HIV has a 10% chance of developing an active case of TB each year. HIV and TB interact in a vicious cycle. HIV compromises the immune system, which allows *M. tuberculosis* to copy itself at higher rates. *M. tuberculosis*, in turn, triggers inflammatory responses and the secretion of transcription factors that directly activate HIV genes. More HIV means more immune destruction—and still more opportunities for *M. tuberculosis* to flourish. A compromised immune system has difficulty containing the infection to the lungs, which means the disease can travel around the body, damaging the brain, the spine, or joints. Diagnosing TB in HIV-infected people is also trickier, so they often do not receive treatment and spread the infection to others. Worldwide, TB is now the leading cause of death in HIV-infected people.

Good news, bad news. Bonnie Davis (*top*) says a new diagnostic, GeneXpert, saved the life of this MDR TB patient at Hlabisa Hospital. Davis's colleague, Tshipuk Martin (*bottom*), says the wards are full because many patients don't adhere to treatments.

CREDITS: MALCOLM LINTON (2)

Joel Ernst, director of infectious diseases at New York University in New York City, says these two bugs likely aid and abet each other by several other mechanisms that scientists have yet to elucidate. "There is so much we don't know about HIV/TB interactions," says Ernst, coauthor of "HIV and Tuberculosis: a Deadly Human Syndemic" in the April 2011 issue of *Clinical Microbiology Reviews*. "I can't think of a better syndemic—or a worse one—than HIV/TB."

The Mbeki administration fueled the syndemic with its pronounced skepticism about the link between HIV and AIDS and its reluctance to introduce antiretrovirals (ARVs). To the astonishment of the medical community inside and outside the country, Mbeki's health minister, Manto Tshabalala-Msimang, went so far as to advocate treatments such as lemons, beetroot, and garlic instead of promoting ARV drugs. "It was very clear that our response to the HIV pandemic in the beginning was the wrong one," says current Health Minister Motsoaledi (see Q&A, p. 902).

HIV prevalence in South Africa nearly doubled between 1998 and today, and many people needlessly died from AIDS. A study done by a team at the Harvard School of Public Health and published in the 1 December 2008 *Journal of Acquired Immune Defi-*

ciency Syndromes calculated that 330,000 lives could have been saved between 2000 and 2005 had Mbeki's administration launched a "feasible" ARV program. TB cases tripled during Mbeki's time in office, and in 2010, nearly half a million South Africans developed active cases—only China and India had a higher incidence—and 60% of them had HIV.

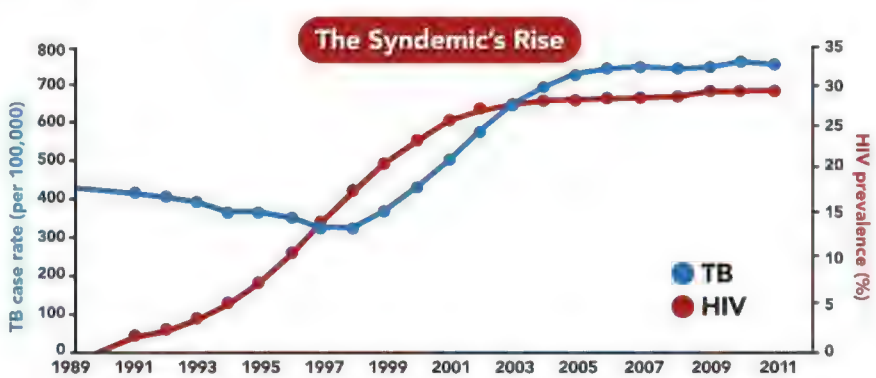
Early warning. Nesri Padayatchi [illegible] drug-resistant TB in kids reflects the [illegible] problem in adults.

A decade ago, South Africa, like many other countries, did little systematic testing for MDR TB—which is more difficult to diagnose than drug-susceptible TB. A 2004 World Health Organization (WHO) report cited only two provinces that had repeatedly surveyed their populations for MDR TB, and the trends suggested that cases were relatively stable. But then in 2006, news of the world's first outbreak of XDR TB in KwaZulu-Natal—53 cases at the same hospital in Tugela Ferry—made clear that the mix of HIV and drug-resistant TB in South Africa had hit a flash point (*Science*, 15 February 2008, p. 894).

Patients with XDR TB don't respond to the first-line drugs or several second-line treatments. Cases of XDR TB had been reported elsewhere before the Tugela Ferry outbreak occurred: Between 2000 and 2004, public health officials had documented 347 sporadic cases worldwide. But no one had seen a cluster of cases, which indicated that an extremely difficult to treat strain of *M. tuberculosis* was likely spreading quickly from person to person. As researchers described at the 2006 International AIDS conference, of the 53 people who had XDR strains at Church of Scotland Hospital in Tugela Ferry, 52 died a median of 16 days after being diagnosed with TB. Although HIV status was not known for nine of the patients, everyone else was infected with the virus.

The Tugela Ferry outbreak set off alarm bells globally, sparking fears of a pandemic of XDR TB and intensifying attempts to find other cases. Last year, WHO said 84 countries had reported cases of similarly difficult to treat *M. tuberculosis* strains; it also estimated that of the world's 650,000 or so cases of MDR TB, 9% were XDR TB.

Within South Africa, the Tugela Ferry outbreak led to extensive studies to determine why it happened and what relation it had to HIV. As it turns out, XDR TB had been bubbling beneath the surface in KwaZulu-Natal for several years, driven by a sharp increase in MDR TB cases that had gone unnoticed. A team led by epidemiologist Kristina Wallengren of the KwaZulu-Natal Research Institute for Tuberculosis and HIV (K-RITH) in Durban reviewed laboratory records for all TB patients who had their *M. tuberculosis* analyzed for drug resistance. In the October 2011 issue of *Emerging Infectious Diseases*, the team reported that between 2001 and 2007, MDR TB cases jumped from 216 to 2799, and XDR cases had a startling 45-fold increase from six to 270.



Feeding the fire. The rise in HIV infections fueled South Africa's explosion of TB cases.

Several of the researchers who first described the Tugela Ferry outbreak published compelling evidence in the 1 January issue of *The Journal of Infectious Diseases* that most people became infected with XDR TB at the Church of Scotland Hospital. The researchers looked at 148 people—a larger group than the one originally described in the outbreak—who had XDR TB between January 2005 and December 2006. Genetic analyses of the *M. tuberculosis* isolates from the patients showed that all but 4% closely matched each other, and 82% of the people had overlapping hospital stays at Church of Scotland. Although 98% of the patients were HIV-infected, only 31% of them were receiving anti-HIV drugs. All had extremely low levels of CD4 white blood cells, leaving them vulnerable to TB.

The study concluded that the Tugela Ferry outbreak resulted from a combination of long hospital stays, poor infection control, delays in the diagnosis of TB drug resistance, and HIV infection. Although the feared pandemic never materialized, the Tugela Ferry outbreak will go down in history for prodding the world to realize that XDR TB had made more inroads across the globe than anyone had imagined. And in South Africa, it also sent a loud message to the country that the response to its HIV/TB syndemic needed a serious overhaul.

Extraordinary measures

On 1 December 2009, World AIDS Day, President Zuma announced sweeping reforms that aimed to prevent and treat infections with both HIV and *M. tuberculosis* more effectively. "We need extraordinary measures to reverse the trends we are seeing in the health profile of our people," Zuma declared.

As Zuma spelled out, the health ministry would launch a "massive campaign" to test 15 million South Africans for HIV. ARVs, until then still restricted to people who had severely damaged immune systems, would become available to HIV-infected people with TB much earlier in the course of the disease. "TB and HIV/AIDS will now be treated under one roof," he promised. Zuma also decided to make ARVs available to all HIV-infected, pregnant women and children under a year of age, regardless of their health status. Fifteen months later, on 24 March 2011—World TB Day—Motsoaledi unveiled a new, equally ambitious, two-pronged approach to diagnosing active cases of tuberculosis. First, the government would ramp up efforts to test the 407,000 families that had a relative with an active case of TB. The second prong called

CREDITS (TOP TO BOTTOM): MALCOLM LINTON; CHERYL BAXTER AND SALIM ABDOOL KARIM, CAPRISA

for a countrywide deployment of a new machine called GeneXpert, which promised to revolutionize the ability to find active TB cases.

The most common TB diagnostic is a sputum "smear" test for *M. tuberculosis* that relies on microscopy; it detects only 60% of active cases and that drops to as low as 35% in HIV-infected patients. Culturing the sputum samples gives the most accurate results, but these tests can take weeks. Drug-sensitivity tests require still more time and costs. GeneXpert, which uses PCR to amplify and detect specific *M. tuberculosis* sequences, in contrast provides a highly accurate diagnosis within 2 hours, and can also detect MDR TB. Although GeneXpert is more expensive than sputum smears or culturing, studies suggest that it will save money in the long run by reducing transmission.

A review article in the 8 December 2012 issue of *The Lancet*, "Health in South Africa: changes and challenges since 2009," shows how these "radical policy changes" have already produced results. Co-authored by CAPRISA head Karim and several other leading South African researchers, the article notes that government funding for TB and HIV/AIDS nearly doubled between 2009–10 and 2010–11. The mass HIV-testing campaign still under way had reached 13 million South Africans by June 2011, up from 2 million. The country quickly developed the largest GeneXpert TB program in the world, performing nearly half a million tests by July 2012. Since April 2010, South Africa has doubled the number of people on ARVs to nearly 2 million, more than any other country. And nearly 400,000 HIV-infected people who have latent *M. tuberculosis* infections take isoniazid, a first-line TB drug that works as a preventive.

Treatment as Prevention, Real World

KWAZULU-NATAL PROVINCE, SOUTH AFRICA—In the summer of 2011, researchers for the first time began seriously discussing the possibility of ending the AIDS epidemic in certain locales. The catalyst for what to many would have seemed like wild-eyed optimism a year earlier was a remarkable finding about the power of antiretroviral (ARV) drugs: If HIV-infected heterosexuals take their ARVs as prescribed, levels of HIV in their bodies fall so low that it reduces the chance of transmission to their long-term partners by 96%. Published in the 11 August 2011 issue of *The New England Journal of Medicine*, the study definitively proved that treatment is prevention. But it had one major weakness: It wasn't "real world," as study participants rigorously adhered to the regimen, taking their medication as directed, which knocked down the virus in their blood to undetectable levels for several years. The cohabiting couples in the study were also more stable than seen in many sub-Saharan communities.

Helping hands. Hlabisa residents participated in a study that proved treatment as prevention works in the real world.

CREDIT: MALCOLM LINTON

Now on page 966, a new study shows that widespread ARV use reduced the spread of HIV in a large population here in KwaZulu-Natal, which has one of the world's most severe HIV/AIDS epidemics—and is as real world as it gets. "We were super excited when we first saw the data because the results were so striking," says epidemiologist Frank Tanser at the Africa Centre for Health and Population Studies in Somkhele who led the study.

Hlabisa, a subdistrict in KwaZulu-Natal, provided an ideal testing ground for a community-level assessment of treatment as prevention. In 2003, the Africa Centre began a population-based survey in Hlabisa of HIV status, sexual behavior, and socioeconomic factors of residents aged 15 and older. The researchers initially hoped their door-to-door survey would help them understand the impact of the HIV/AIDS epidemic in the region and its demographic contours. The survey coincided with a massive rollout of ARVs in KwaZulu-Natal, which the center helped coordinate, so this gave them a new opportunity to look at treatment as prevention. To date, the center's efforts have helped start more than 20,000 people on anti-HIV drugs.

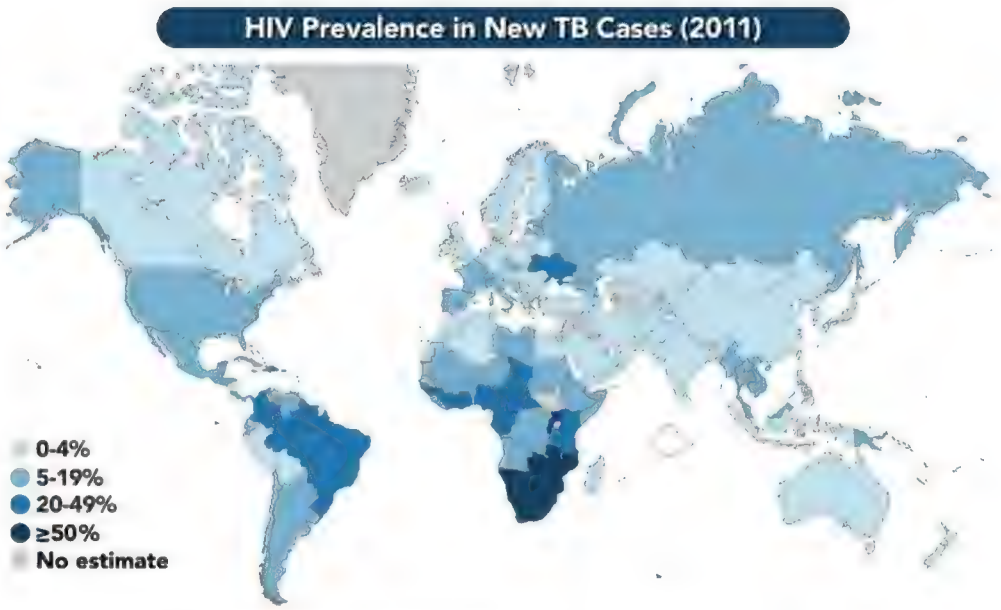
The Africa Centre team identified 16,667 people who were uninfected with HIV in 2003 and agreed to have repeated tests for the AIDS virus each year. Between 2004 and 2011, 1413 of these people became infected.

Researchers at the Africa Centre had previously shown that sexual partnerships in Hlabisa have strong geographical links. In a study published by Tanser and colleagues in the 16 July 2011 issue of *The Lancet*, 67% of the participants reported having sex over a 5-year period with someone in their small Zulu community known as an *isigodi*. So for the new study, the investigators created a circle around each uninfected person's home that had a radius of 3 kilometers and then analyzed the proportion of HIV-infected people receiving ARVs in that small area. If treatment worked as prevention on a community level, then, theoretically, there should be fewer new infections in neighborhoods where a higher percentage of infected people were on ARVs.

During the study, ARV coverage jumped from less than 10% of the infected population to 37%. Tanser's group found that the new infection rate was lowest in places that had the highest ARV usage: If ARV coverage was 30% to 40%, the people in those communities had 38% fewer new infections when compared with those in locales that had ARV coverage of less than 10%.

What's more, the study showed that not all infected people, and certainly not 100%, need to be on treatment to see an effect. "Once you get to 30% coverage, you get a steep decline in new infections," Tanser says. So in the end, this is one of the rare examples in which the real world is a more forgiving place than many people imagined. **–J. C.**

Tangible benefits have begun to surface, too. The proportion of deaths related to HIV/AIDS has dropped from 52.3% of the total in 2006 to 43.6% in 2011. (Death rates from TB, excluding HIV, have remained relatively constant.) As a study in this issue reports, life expectancy in KwaZulu-Natal went from 49.2 in 2003—the year before the government began providing ARVs—to 60.5 in 2011 (see p. 961). Only 2.7% of babies born to HIV-infected women in 2011 were infected after 6 weeks, down from a high of 20% to 30% a decade earlier. "Undeniably, much remains to be done," the report concludes. "However, for the first time in two decades, this progress instills a basis for hope."



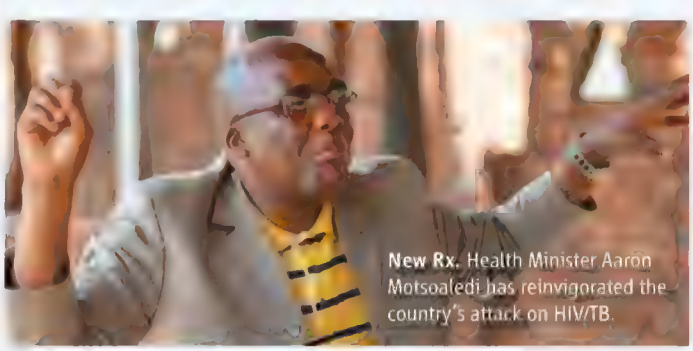
Hardest hit. The HIV/TB syndemic has walloped southern Africa.

Several other recent developments illustrate how far South Africa has come in comparison to its neighbors. A study in this week's issue of *Science* shows that widespread use of ARVs in KwaZulu-Natal reduced the spread of HIV in the community (see p. 966)—the first demonstration outside a clinical trial that treatment works as prevention in a sub-Saharan African setting. The country also has become a world leader in HIV/TB research, and now boasts its own world-class institute devoted to the study of these married diseases. With $40 million from the Howard Hughes Medical Institute, in October 2012 K-RITH opened a swank 4000-square-meter, 8-story building that features several high-tech biosafety spaces with negative air pressure that can handle dangerous pathogens, state-of-the-art DNA sequencing machines, and the fastest Internet connection available in the country. "We're trying to put the technology where the problem is," says K-RITH's director, William Bishai, who maintains a lab at Johns Hopkins University in Baltimore, Maryland.

In the next 3 years, the government says in its National Strategic Plan for these diseases that it hopes to reduce new HIV and TB infections by 50%, which it will do

Pulling South Africa Back From the HIV/AIDS Brink

When Aaron Motsoaledi, the South African minister of health, spoke in Vienna at the 18th International AIDS Conference in 2010, he was largely unknown to that community, which had come to see his government as a pariah in the fight against HIV/AIDS. At the Vienna meeting, Motsoaledi, a medical doctor who was appointed to the health ministry in May 2009 shortly after President Jacob Zuma took office, announced to the audience that his country had done an about-face. Today, many of the goals that Motsoaledi spelled out in his Vienna talk have been met, as the country has aggressively expanded testing and treatment for both HIV and TB (see main story, p. 898).



New Rx. Health Minister Aaron Motsoaledi has reinvigorated the country's attack on HIV/TB.

Motsoaledi met with *Science* in September 2012 in the courtyard of an aging but elegant conference center atop a hill in Pretoria. The center was hosting President Zuma and other leaders attending an African National Congress (ANC) meeting. An executive member of ANC who during the antiapartheid struggle once worked with an underground, armed wing, Motsoaledi was candid and refreshingly blunt, never sidestepping controversial issues or requesting that his thoughts be off the record. **–J. C.**

Q: The International AIDS Conference in Durban, South Africa, in 2000 was a turning point in the epidemic, as the world recognized that treatments saving lives in rich countries should be available to the poor. Yet your president at the time, Thabo Mbeki, was questioning whether HIV caused AIDS. What did you think of your country then?

A.M.: I was very confused and extremely disappointed. It was one of the lowest moments in the country. I'm a trained doctor, so I was also aware and worried, but I was not [working] within the health sector.

Q: What did you do when you came in as health minister?

A.M.: Now that I was given authority, it didn't need rocket science for me to know exactly what it is that I need to do, and I went out to do it. I was not the only one. We had a new chairperson of the South African National AIDS Council, the Deputy President Kgalema Motlanthe, who had similar beliefs. On top of everything else we had President Zuma, who also complained this was a wrong policy, and now that he was in charge he cannot allow it. All I had to do was to show the main figues, exactly what the scourge has done.

CREDITS (TOP TO BOTTOM): WORLD HEALTH ORGANIZATION, *GLOBAL TUBERCULOSIS REPORT 2012*; MALCOLM LINTON

by ensuring that everyone in the country receives an annual HIV test and an exam for TB symptoms. It also aims to increase ARV coverage to 80% and drive mother-to-child transmission below 2%. As the strategic plan notes, achieving these aspirations depends on substantial funding increases that will have to come mainly from government coffers. K-RITH's Bishai also soon hopes to have a dozen leading HIV and TB researchers from around the world investigating how the two diseases interact. "It's a black box," says Bishai, whose own work focuses on the genetics of drug resistance. "And the fields have been siloed. There's almost no co-authorship."

Despite these steps forward, Hlabisa Hospital, located atop rolling hills that abut one of several game parks in KwaZulu-Natal, reveals the staggering challenges that exist today. Tshipuk Martin, the doctor in charge of HIV/AIDS at the 275-bed hospital, in September 2012 led a tour of the crowded female adult ward, where some patients slept on mattresses on the floor. "Up to 80% are here for HIV, and all have complications of opportunistic infections," Martin says. "Almost all patients are coming in on ARVS, but they're coming in very sick and the problem is adherence" to drug regimens.

Success story. Drugs cured Nomusa Manqele's TB and have controlled her HIV infection. "People have trouble with adherence because they're not accepting their status," Manqele says. "I tell them to look at me. I'm fine. I'm healthy. I'm proud of my status. I told myself the stigma, I'd get rid of it."

Improved diagnostics and better use of existing drugs alone will not curb the HIV/TB syndemic. That will require even wider use of ARVs and existing antibiotics, as well as new treatments that cure drug-sensitive *M. tuberculosis* more quickly and have more power against resistant strains. (The United States just approved a new drug, bedaquiline, that combats MDR TB [*Science*, 11 January, p. 130].) Effective vaccines against both diseases, which remain elusive, would make huge inroads, too. But South Africa clearly has more tools—and more political will—to confront HIV/AIDS and TB than ever before.

Then again, meaningful change takes time, and it ran out for Sandile on 8 October. Despite specialty care at King George V, the tag-team drubbing from HIV and *M. tuberculosis* finally proved too much for the young boy's body, leading to irreversible kidney damage and sepsis. During his entire 6 months of his hospitalization, he never had a visitor. One week passed before relatives arrived from rural KwaZulu-Natal to retrieve his body, which they took back to their village for a traditional funeral.

–JON COHEN

Q: What figures did you show?

A.M.: In this country, it's not because of a lack of research and knowledge that a wrong policy was followed. HIV/AIDS is a sort of animal, and many different groups had different parts of this animal. Nobody knew exactly what the animal looked like. I realized that if we brought these parts together, we'd start seeing whether we're dealing with a big animal or a mouse—whether it's a monster or something you can just kick and say, ay, this is nothing. I went back to 2000 and collected every piece of evidence about HIV in South Africa from internal and external organizations, and I put it together in a PowerPoint. And this started showing a clear picture of what's happening all over the country. Before I did this, some very prominent members of society were saying, "Minister, are you not exaggerating this HIV/AIDS problem? Don't we have diabetes, high blood pressure, cancer? We never hear you mention these things—every time you open your mouth it's HIV/AIDS. Don't you understand the danger of becoming the minister of HIV/AIDS instead of the minister of health?" Now these graphs balanced things.

Q: Who saw this PowerPoint?

A.M.: The first person I targeted was the president. It was October 2009, and it was half past 11 in the night when he gave me an appointment. I asked for 45 minutes with him. He was shocked. He said, "I've been knowing what we were doing was wrong but no one ever showed me in numbers and figures." He said, "Tomorrow, you are going to present this to the Cabinet." And I did. And some members of the Cabinet also said, "Oh, these figures and numbers, ay, something out of this world. We never saw them." They might have seen them somewhere but in different parts, and putting them together, they said, "This is no ordinary animal—it's a monster." And that's why everything started.

Q: Some developing countries think South Africa has made great strides because it has more resources. Is that true?

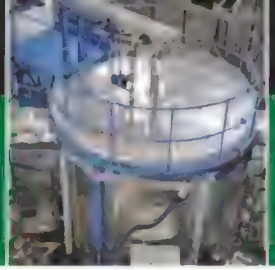
A.M.: I don't necessarily think so. It was more of a political will. I didn't only show the PowerPoint to the Cabinet, I showed it to all the ministers, including the minister of finance. So all of the ministers are united. The public turning point happened on World AIDS Day, first of December 2009, when the president shared the platform with the executive director of UNAIDS [the Joint United Nations Programme on HIV/AIDS], Mr. Michel Sidibé, who was also very helpful. The president made far-reaching announcements about how to deal with HIV/AIDS. The budget was already completed for the full year, but he went to the minister of finance and said, You have to change the budget to suit this announcement. Other ministers had to accept budget cuts.

Q: What frustrates you and makes you think the country is not doing what you hope it could do or should do?

A.M.: You need highly skilled people to deal with this problem. We are struggling with that. There's a massive shortage of health workers around the whole world, but in sub-Saharan Africa especially. Secondly, the burden of disease is very high. And because of our wrong approach in the beginning, the program started very late, and it means we're lagging behind. We lost a decade, and that makes it very difficult to catch up.

CREDIT: MALCOLM LINTON

Old ways worth following 907



To haze and haze not 911

LETTERS

edited by Jennifer Sills

Old Trees: Extraction, Conservation Can Coexist

BECAUSE LARGE OLD TREES ARE ESSENTIAL FOR FOREST ECOSYSTEM INTEGRITY AND BIODIVERSITY, timber extraction in managed forests should preferentially be concentrated where large old trees are least likely to develop ("Global decline in large old trees," D. B. Lindenmayer *et al.*, Perspectives, 7 December 2012, p. 1305). However, timber extraction and the conservation of large old trees are not necessarily mutually exclusive.

Current forest policy and management practices in Flanders, Belgium, aim to convert even-aged stands (areas in which trees are all the same age) to stands with trees of varying ages in an effort to increase forest ecosystem stability and resilience and to allow trees to grow old. As part of their ecologically sustainable forest management, public forest managers have adopted a large-tree retention approach [see also (*1*, *2*)]. Tree islands within stands managed for production of high-quality timber are reserved for conservation, and trees within these islands will never be extracted. Large old trees of commercially valuable species that have grown beyond the commercially optimal dimensions will not be logged either. And no tree beyond a threshold diameter [currently set at dbh (diameter at breast height) of more than 102 cm] will ever be logged. The strip-shelterwood system (in which trees are cut in linear strips and surrounding trees are given time to grow old) and the coppice-with-standards system (in which some trees are left to grow while others around them are cut) are two examples of forest management that allows the combination of sustainable forest exploitation and conservation of large old trees.

RAF AERTS

Division Forest, Nature and Landscape, University of Leuven, Celestijnenlaan 200E-2411, BE-3001 Leuven, Belgium. E-mail: raf.aerts@biw.kuleuven.be

References

1. D. B. Lindenmayer, W. F. Laurance, *Biol. Conserv.* **151**, 11 (2012).
2. D. B. Lindenmayer *et al.*, *Conserv. Lett.* **5**, 421 (2012).

Old Trees: Cultural Value

AS D. B. LINDENMAYER *ET AL.* POINT OUT ("Global decline in large old trees," Perspectives, 7 December 2012, p. 1305), large old trees play a key ecological role in many different environments, and their observed decline may have disastrous consequences for biodiversity and ecosystem integrity. However, the value of large trees as part of our cultural heritage, often neglected in conservation, may be essential for addressing the problem of their global decline.

Old trees span several human generations and thereby constitute a living link between them. Cultural, religious, spiritual, and symbolic values of the large trees, as well as indigenous communities' reliance on services provided by them, provide a firm foundation for practical conservation. In many cases, the general society's preferences concerning large trees may coincide with the broad conservation interest (*1*).

Moreover, large trees wake emotions, appeal to our aesthetic sentiments, and are often perceived as important landmarks. In this sense, they may be perfect flagship elements of the conservation strategies. For example, a concern for big trees in the beginning of the 20th century led to action that safeguarded the future of giant sequoias, perceived as "natural temples" (*2*). When considering a multitude of large trees' virtues, it seems that the value of these iconic organisms should be more broadly recognized by the conservation community, as it may support their conservation goals.

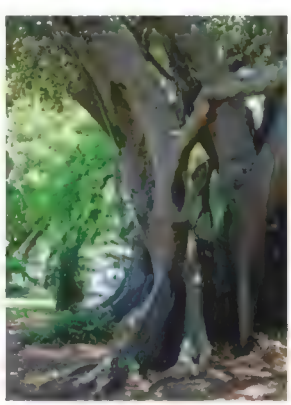
MALGORZATA BLICHARSKA[1]* AND GRZEGORZ MIKUSINSKI[2]

[1]Department of Aquatic Sciences and Assessment, Swedish University of Agricultural Sciences, 750 07 Uppsala, Sweden. [2]School for Forest Management, Swedish University of Agricultural Sciences, 739 21 Skinnskatteberg, Sweden.

*To whom correspondence should be addressed. E-mail: Malgorzata.Blicharska@slu.se

References

1. A. M. Lykke, *J. Environ. Manage.* **59**, 107 (2000).
2. *Science* **54**, 43 (1921).

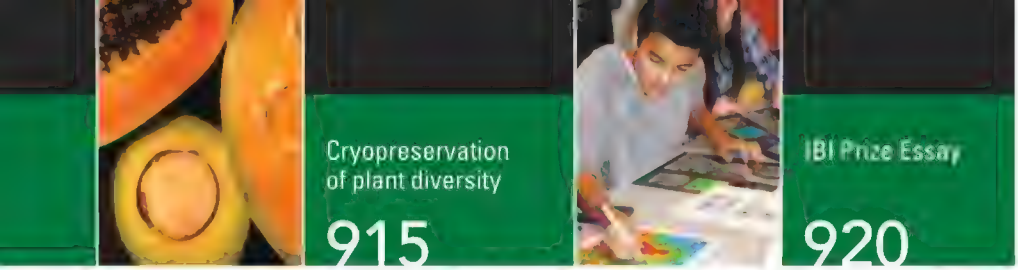


CREDIT: GRZEGORZ MIKUSINSKI

Old Trees: Large and Small

D. B. LINDENMAYER *ET AL.* ("Global decline in large old trees," Perspectives, 7 December 2012, p. 1305), report a global decline in large old trees and show that climate change and human disturbance are reducing the abundance of these ecologically important organisms. Such framing of the problem leads to conflation of two issues: Old trees and large trees are not synonymous.

While the term "old growth" brings to mind the iconic sequoias described by the authors, stunted and slow-growing forests in extreme environments also play important ecological roles, and declines in small old trees are of increasing concern. Landscapes with small old trees (often surviving at the edge of their ranges) function as genetic repositories and population refugia, which play a critical role in the long-term persistence of forest ecosystems (*1*, *2*).



In southern Brazilian Atlantic forests, where highly diverse isolated populations can persist over several millennia, small old trees have shown recent growth decline (*3*). Here, at the southernmost limit of tropical forest distribution, trees more than 40 cm in diameter are rare, and landscapes dominated by small trees provide essential services, but very little of the original forest cover is protected by nature reserves. A hemisphere away, in northern Ontario, Canada, live some of the largest undisturbed old-growth pine forests in North America (*4*). Recent work suggests that here, too, the growth of small old trees (typically less than 50 cm in diameter, from 150 to more than 300 years) has been declining (*5*). As climatic changes and human demands increase, widespread growth declines and increased mortality threaten global forests (*6*–*8*). The struggles of large old trees are important, but they are by no means unique.

JACOB CECILE,[1] LUCAS R. SILVA,[2] MADHUR ANAND[1]*

[1]Global Ecological Change and Sustainability Laboratory, School of Environmental Sciences, University of Guelph, Guelph, ON, N1G 2W1, Canada. [2]Department of Land, Air, and Water Resources, University of California Davis, Davis, CA 95616, USA.

*To whom correspondence should be addressed. E-mail: manand@uoguelph.ca

References

1. J. S. McLachlan, J. S. Clark, P. S. Manos, *Ecology* **86**, 2088 (2005).
2. L. C. R. Silva, M. Anand, *Ecosystems* **14**, 1354 (2011).
3. L. C. R. Silva, M. Anand, J. M. Oliveira, V. D. Pillar, *Glob. Change Biol.* **15**, 2387 (2009).
4. M. D. Leithead, M. Anand, L. C. R. Silva, *Oecologia* **164**, 1095 (2010).
5. L. C. R. Silva, M. Anand, M. D. Leithead, *PLoS ONE* **5**, 7 (2010).
6. L. Andreu-Hayles *et al.*, *Global Change Biol.* **17**, 2095 (2011).
7. P. J. van Mantgem *et al.*, *Science* **323**, 521 (2009).
8. B. Choat *et al.*, *Nature* **491**, 752 (2012).

Response

CECILE *ET AL.* TAKE ISSUE WITH OUR RECENT Perspective on the rapid global decline of large old trees by asserting that "large" and "old" are not synonymous. Of course, some ancient trees are indeed short in stature [e.g., (*1*)]. Nonetheless, many of the world's largest trees are also old (more than 500 to 1000 years) (*2*, *3*), so it is correct to highlight this reality.

Cecile *et al.* further argue that small old trees are declining in some ecosystems, and we agree that this is a matter of concern. However, large old trees are particularly vulnerable to disturbances such as insect attack, pathogens, drought, and fire in some environments (*4*–*6*). Large old trees also play a range of key roles (e.g., as wildlife habitat and for carbon storage) that are not played by small old trees. Unlike large old trees, small old trees are rarely deliberately removed during agricultural intensification (*7*) or because of safety concerns in urban environments (*4*). In addition, small old trees are less likely to be targeted for logging [although they may be specifically targeted in some restoration projects in western North America because of prohibitions on removal of larger trees (*8*)]. The particular risks faced by large old trees, combined with their unique ecological roles, mean that their management and conservation warrant special mention.

DAVID B. LINDENMAYER,[1]* WILLIAM F. LAURANCE,[2] JERRY F. FRANKLIN[3]

[1]Fenner School of Environment and Society, The Australian National University, Canberra, ACT 0200, Australia. [2]Centre for Tropical Environmental and Sustainability Science (TESS), and School of Marine and Tropical Biology, James Cook University, Cairns, QLD 4878, Australia. [3]School of Environmental and Forest Science, University of Washington, Seattle, WA 98195, USA.

*To whom correspondence should be addressed. E-mail: david.lindenmayer@anu.edu.au

References

1. Umeå University, Press Release, "World's oldest living tree discovered in Sweden" (16 April 2008); http://info.adm.umu.se/NYHETER/PressmeddelandeEng.aspx?id=3061.
2. J. Chambers, N. Higuchi, J. P. Schimel, *Nature* **391**, 135 (1998).
3. W. F. Laurance *et al.*, *Forest Ecol. Manage.* **190**, 131 (2004).
4. B. J. Palik, M. E. Ostry, R. C. Venette, E. Abdela, *Forest Ecol. Manage.* **261**, 128 (2011).
5. M. Simard, E. N. Powell, K. F. Raffa, M. G. Turner, *Glob. Ecol. Biogeogr.* **21**, 556 (2012).
6. D. C. Nepstad, I. M. Tohver, D. Ray, P. Moutinho, G. Cardinot, *Ecology* **88**, 2259 (2007).
7. M. Maron, J. A. Fitzsimons, *Biol. Conserv.* **135**, 587 (2007).
8. J. F. Franklin, K. N. Johnson, *J. Forestry* **110**, 429 (2012).

Letters to the Editor

Letters (~300 words) discuss material published in *Science* in the past 3 months or matters of general interest. Letters are not acknowledged upon receipt. Whether published in full or in part, Letters are subject to editing for clarity and space. Letters submitted, published, or posted elsewhere, in print or online, will be disqualified. To submit a Letter, go to www.submit2science.org.

China's Little Emperors Show Signs of Success

IN THE REPORT "LITTLE EMPERORS: BEHAVIOR impacts of China's one-child policy" (22 February, p. 953; published online 10 January), L. Cameron *et al.* suggest that being an only child as a result of China's one-child policy (OCP) might carry negative implications for success and well-being. We caution against overgeneralization from the economic experiments and personality surveys conducted in this study.

In longitudinal studies among the first post-OCP cohort—the same cohort as that studied by Cameron *et al.*—some singletons showed more behavioral problems and less independence in childhood. Yet, by adolescence, differences in behavioral problems disappeared and independence levels reversed (*1*, *2*). In contrast to the experimental context, more pro-social behaviors among singletons than non-singletons were found in community samples (*2*). In education, singletons performed as well as or better than their peers in verbal and math skills and showed better school adjustment and lower levels of anxiety, but they demonstrated inferior testing and study methods (*3*, *4*). Once they became parents, this cohort of singletons showed no difference in marriage adjustment compared with their counterparts, and in fact demonstrated higher levels of life satisfaction and higher intergenerational family fusion (*5*). Contrary to the hypothesis that maladjustment might be magnified in later post-OCP cohorts, singletons born in the 1990s have shown equal or superior interpersonal relationships with peers, teachers, and family members compared with non-singletons (*6*).

All this is not to deny legitimate concern over the OCP, only-child status, or the combination of both. Yet, against expectation and stereotype, research has shown that as "little emperors" transition into adulthood, their well-being and performance are comparable with, if not superior to, those with one or more siblings. In such complex systems as that of human psychology and behavior, one must move beyond linear notions of causality. Circular processes of self-correction at the individual, family, and social levels often provide surprising compensatory responses to initial conditions.

XUDONG ZHAO,[1] XIQUAN MA,[1] YUHONG YAO,[2] CHONGHUA WAN,[3] EMILY NG[4]*

[1]Shanghai East Hospital; Department of Psychosomatic Medicine, School of Medicine at Tongji University, Shanghai, China. [2]Psychological Counseling and Education Center, Tongji University, Shanghai, China. [3]School of Humanities and Management, Guangdong Medical College, Dongguan, China. [4]Department of Anthropology, University of California, Berkeley, Berkeley, CA 94720, USA.

*To whom correspondence should be addressed. E-mail: emily.ng@berkeley.edu

References

1. M. Wang *et al.*, *World Cultural Psychiatry Res. Rev.* **2**, 118 (2007).
2. X. Feng, *Zhongguo Shehui Kexue* (Social Sciences in China) **13**, 118 (2000).
3. T. Falbo, D. L. Poston, *Child Dev.* **64**, 18 (1993).
4. Z. Li, S. Wu, X. Zhang, *Qingnian Yanjiu* (Youth Research) **4**, 31 (1998).
5. X. Ma, F. Yin, Y. Yao, X. Zhao, *Zhongguo Xinli Weisheng Zazhi* (Chinese Mental Health Journal) **2**, 118 (2012).
6. Y. Yao, doctoral dissertation, School of Medicine at Tongji University (2011).

CORRECTIONS AND CLARIFICATIONS

This Week in *Science*: "Proton still too small" (25 January, p. 371). The study examined muonic hydrogen, not muonium. Muonic hydrogen is the name given to a system made from a proton and a negative muon. Muonium is the exotic atom made from a positive muon and a (negative) electron. The HTML and PDF versions online have been corrected.

News Focus: "The children's study: Unmet promises" by J. Kaiser (11 January, p. 133). The January 2013 workshop to review the study plan described in the text and timeline was sponsored by the National Academy of Sciences, not the Institute of Medicine. The HTML and PDF versions online have been corrected.

Letters: "NextGen speaks" (4 January, p. 30). Ali Jawaid's essay referred to Pakistan, not Switzerland. Jiang Zhao is at Beihang University, not Beijing University. Guilherme Martins Santos's address should be Laboratory of Molecular Pharmacology, Department of Pharmaceutical Sciences, Faculty of Health Sciences, University of Brasília, Brazil, CEP 70910-900, Brazil. In the online-only essays, Homare Yamahachi's essay referred to Greece, not Norway. These changes have been made in the HTML and PDF versions online.

Reviews: "Conversion of wastes into bioelectricity and chemicals by using microbial electrochemical technologies," by B. E. Logan and K. Rabaey (10 August 2012, p. 686). In the Fig. 2 legend, the symbols should have been identified as follows: "There is roughly an inverse relationship between the value of these products (circles) and the current densities (triangles)...." On p. 689, "4.5 a.m.$^{-2}$" should have been written as "4.5 A m^{-2}." The authors also wish to thank T. Lacoere and J. Desloover for assistance in preparing the draft figures and funding from the Commonwealth Scientific and Industrial Research Organization Flagship cluster "Biotechnological solutions to Australia's transport energy and greenhouse gas challenges." The HTML version online has been corrected.

TECHNICAL COMMENT ABSTRACTS

Comment on "Bilaterian Burrows and Grazing Behavior at >585 Million Years Ago"

Claudio Gaucher, Daniel G. Poiré, Jorge Bossi, Leda Sánchez Bettucci, Ángeles Beri

Pecoits *et al.* (Reports, 29 June 2012, p. 1693) describe bilaterian trace fossils and assign them an Ediacaran age based on the age of a granite interpreted as intrusive. We argue that the granite is not intrusive but in fact represents the basement of the sedimentary succession. Moreover, we show that identical trace fossils occur in nearby Carboniferous-Permian glacigenic rocks.

Full text at http://dx.doi.org/10.1126/science.1230339

Response to Comment on "Bilaterian Burrows and Grazing Behavior at >585 Million Years Ago"

Ernesto Pecoits, Kurt O. Konhauser, Natalie R. Aubet, Larry M. Heaman, Gerardo Veroslavsky, Richard Stern, Murray K. Gingras

Gaucher *et al.* suggest that their field observations and petrographic analysis of one thin section do not support an Ediacaran age for the trace fossils–bearing strata of the Tacuarí Formation. We have strengthened our conclusion of an Ediacaran age for the Tacuarí Formation based on reassessment of new and previously presented field and petrographic evidence.

Full text at http://dx.doi.org/10.1126/science.1230677

CREDIT: WILLIAM LEAMAN/ALAMY

ANTHROPOLOGY

Lessons from Our Elders

W. G. Runciman

The World Until Yesterday
What Can We Learn from Traditional Societies?
by Jared Diamond
Viking, New York, 2012. 543 pp. $36, C$38. ISBN 9780670024810. Allen Lane, London. £20. ISBN 9780713998986.

Within a single lifetime, New Guinea highlanders who until the 1930s knew nothing of their fellow human beings in the modern world were catapulted across ten millennia of cultural and social evolution. Jared Diamond's natural- and social-scientific interests and many years of exploration and research in New Guinea make him well qualified to expound what he sees as the implications both for them and for us. In *The World Until Yesterday*, he also draws on the ethnography of a number of other well-documented societies that he labels "traditional" as opposed to "state," including the !Kung, Ache, Siriono, Yanomamo, Shoshone, and Inuit. But it is the people of New Guinea whom he has principally in mind.

Traditional transport. Vertically erect and facing forward, the carried child (here a Pumé Indian baby from Venezuela) sees the same field of view as the caregiver (an older sister).

CREDIT: RUSSELL D. GREAVES AND KAREN KRAMER

Admirers of *Guns, Germs, and Steel* (*1*) may be surprised to find that in this book, Diamond has almost nothing to say about the range of state societies that has evolved since the transition from hunting and foraging to agriculture. Largely skipping the development of sedentism, markets, armies, churches, and bureaucracies and the concentration of increasingly expanding populations in increasingly large communities, he goes more or less straight from his traditional societies to the present-day United States. In a chapter on religion, he makes in passing the claim that the book is "devoted to exploring the whole spectrum of human societies, from small-scale or ancient to populous or modern." But it does no such thing, beyond summarily differentiating the functions of religion in bands and tribes by 5000 BCE, rising chiefdoms and states from 5000 to 1000 BCE, European religious states at around 1600 CE, and the secular states of today.

The sociological taxonomy on which Diamond relies is Elman Service's limited and by now heavily qualified distinction among bands, tribes, chiefdoms, and states (*2*). Much of the evidence for Diamond's conclusions is frankly anecdotal. He is not concerned with either the underlying process or the detailed mechanisms of cultural and social evolution, but only with the traits of traditional societies that he thinks could be imitated by state societies to the latter's benefit.

Diamond's recommendations fall under seven headings: the settlement of disputes; the bringing up of children; the treatment of the old; and attitudes toward risk, religion, language, and physical health. He has no romantic illusions about hunter-gatherer lifestyles and well understands why New Guineans embrace the benefits that American society has to offer them. But he is convinced that we have lessons to learn from them as well as the other way round.

Different readers will have different opinions about both the originality (or, some will say, banality) of the recommendations and about the prospect (or, some will say, unlikelihood) of their being put into effect. Diet is the most straightforward: the reasons why excessive intake of sugar and salt both damages health and shortens life expectancy are well understood, and the means of combating it are at hand. It is less easy to see how Americans will be persuaded that all children should be multilingual. Nor can one easily identify the practical implications of the injunctions that we should "be honest with ourselves about what religion does or might mean specifically to us" or that in assessing the dangers inherent in our lifestyles we should adopt "New Guinea–style constructive paranoia selectively." (One of his chosen examples is the failure of Harvard's investment managers in the run-up to the 2008–09 financial meltdown to "have followed the [risk-management] strategy practised by so many peasant farmers.") Perhaps more could be done to enrich the lives of elderly Americans by keeping them in closer contact with children and grandchildren, as "traditional" families have no choice but to do. But how likely is it that American parents will take to practicing on-demand nursing and late weaning, let alone stop buying their children video games? And whereas there is much to be said for mediation in both interpersonal and wider disputes and conflicts, how much scope for it is there in a society whose legal system is among the most adversarial in the world?

Although a professor of geography, Diamond twice describes himself as an evolutionary biologist. When, therefore, he ventures into comparative sociology, he cannot but be aware that by comparison with biology, it is a science still in its infancy. Sociologists have nothing approaching a Linnaean classification of societies with a Darwinian rationale. Natural experiments where the relevant variables are adequately controlled are few and far between. As Diamond remarks à propos the unimpeachable generalization that all human societies practice both violence and cooperation, "which trait appears to predominate depends on the circumstances." But the workings of cultural and social variation and selection are still far too little understood for it to be possible to specify the circumstances under which one trait is more likely than another to be reproduced, whether by homologous descent, lateral diffusion, or convergent evolution.

What we (sociologists, that is) do, however, know is the extent to which the trajectories of cultural and social evolution in all human societies are path-dependent. Whether in cultural selection (where information affecting phenotype is acquired by imitation or learning) or in social selection (where it is imposed on phenotypic behavior by rules encoded in institutional sanctions), mutations adaptive in one environment may be either exaptive or maladaptive in another. Only hindsight discloses which. But this sociologist's guess is that Diamond's recommendations to his fellow Americans have not much chance of being put into effect.

The reviewer is at Trinity College, Cambridge CB2 1TQ, UK. E-mail: wgr@wgrunciman.u-net.com

References
1. J. Diamond, *Guns, Germs, and Steel: The Fates of Human Societies* (Norton, New York, 1997).
2. E. R. Service, *Primitive Social Organization: An Evolutionary Perspective* (Random House, New York, 1962).

10.1126/science.1234201

ECONOMICS

A Journey Through 200 Years of Models

Caterina Marchionni

To many observers, economics appears dry, impersonal, and far removed from actual economic concerns. With its heroes, caricatures, recipes, analogies, narratives, and much more, *The World in the Model* offers a very different picture. Mary Morgan's book is a "travel guide" through two centuries of economic modeling, including François Quesnay's *Tableau Économique*, David Ricardo's model farm, Alfred Marshall's diagrams of supply and demand, John Hicks's IS/LM model of the Keynesian system, and the prisoner's dilemma game. Her guide makes the reader appreciate economics much as we come to appreciate a city we had only heard about. We discover its beauty by admiring its landmarks and learning who built them, how they were built, what they represented in the past, and what they mean today.

The World in the Model
How Economists Work and Think
by Mary S. Morgan
Cambridge University Press, Cambridge, 2012. 439 pp. $125, £75. ISBN 9781107002975. Paper, $39.99, £24.99. ISBN 9780521176194.

Not only a refreshing history of economic modeling, the book also presents an insightful philosophy of economic modeling. Morgan (a historian at the London School of Economics and the University of Amsterdam) conceives of economic models as little worlds that represent aspects of the economy and can be analyzed and manipulated in different ways. To be suitable working objects for economists, they must be small and simplified (so that they can be analyzed). Nonetheless, their format and content must be rich enough to permit productive manipulations. Economic models are "objects to enquire into," manipulated and analyzed in order to understand their properties and behavior (thereby helping economists probe theories and intuitions). They are also "objects to enquire with," employed in order to gain insights about the actual world.

According to Morgan, the question of when results from a model can be transferred to the world does not concern the truth of the model vis-à-vis its target. Instead the question concerns the validity of the inferences from the model to the world. The problem economists face is analogous to the external validity problem experimenters face.

Morgan argues that narratives provide the connection between economic models and the world. Narratives are "stories" that a model allows economists (working with and through it) to assemble. (For example, narratives can relate how changes in certain elements of the model causally influence others.) They also serve to indicate possible points of correspondence between elements of the model and aspects of the world. As in many experimental contexts, Morgan notes, model-based inferences are informally made. Yet manipulating a model is not like manipulating some portion of the natural or social world as it occurs in laboratory experiments. And this is why, from an epistemic point of view, in spite of their similarity, model-based inferences are less powerful than experimental ones.

Morgan's account comes to full life in her many detailed examinations of particular cases of both model making and model use. For example, she offers vivid accounts of the development of the Edgeworth box diagram, the Newlyn-Phillips machine, and Orcutt's microsimulation—three models (a diagram, a hydraulic machine, and a computer-based simulation) that differ greatly in format but whose building processes, she shows, all involved elements of cognition and creativity.

Hydraulic model. Bill Phillips with the Newlyn-Phillips macroeconomics machine.

CREDIT: IMAGELIBRARY 6/LONDON SCHOOL OF ECONOMICS AND POLITICAL SCIENCE

Morgan also reconstructs the history of that key figure in economics, *Homo economicus*, from its ancestry in Thomas Malthus and John Stuart Mill to today's rational economic man. She advocates viewing the rational economic man as a caricature rather than a realistic portrait. Caricatures exaggerate certain features, and distort others, of their subject. To the knowing, however, good caricatures recognizably represent their subject. The same holds for the economic man, Morgan argues, which makes it a workable model for economists. For example, Morgan shows that it is the right character for the prisoner's dilemma and other game-theoretic models, in which it acts according to the kinds of situation each model represents.

Although throughout the book Morgan exhibits a sympathetic, almost affectionate, attitude toward the models she examines, she also draws attention to their limitations. Contrary to the views of some critics of economics, however, she doesn't see the culprit to be formal modeling nor the way in which modeling works in the discipline. For her, assessing economic models is mainly a local matter: a judgment about particular models and their usefulness for specific purposes. This important point naturally follows from her view of economic models as working objects. Yet, it is perhaps equally important to investigate which systematic biases affect the ways in which models are built and used in economics [c.f., (*1*)]. Such biases might help explain why economic modeling does not work in some contexts, even if it works well for the kinds of questions economists ask. At the same time, understanding the success and failure of particular models contributes to the general task of identifying the characteristic biases of the economic approach to modeling.

The skillful combination of history and philosophy of science makes *The World in the Model* a must-read for historians and philosophers of economics and of science. Many others, including economists, will enjoy Morgan's tour of economic modeling. All should come away with a better understanding of the small worlds economists work with and a more balanced assessment of their epistemic value.

References

1. W. C. Wimsatt, *Re-Engineering Philosophy for Limited Beings: Piecewise Approximations to Reality* (Harvard Univ. Press, Cambridge, MA, 2007); reviewed in (*2*).
2. K. Sterelny, *Science* **321**, 344 (2008).

The reviewer is at the Finnish Centre of Excellence in the Philosophy of the Social Sciences, Department of Political and Economic Studies, University of Helsinki, FI-00170 Helsinki, Finland. E-mail: caterina.marchionni@helsinki.fi

10.1126/science.1232431

INFECTIOUS DISEASE

The Stability of Malaria Elimination

Eradication may not be necessary before countries can eliminate, scale back control, and rely on health systems.

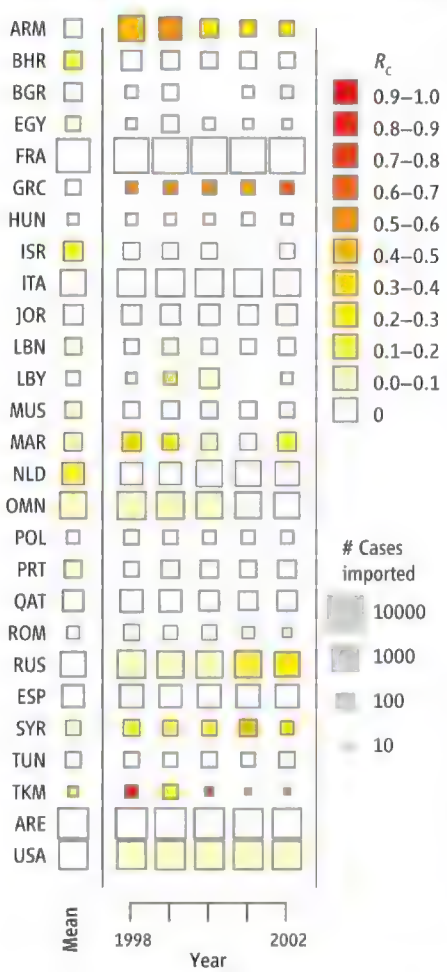
C. Chiyaka,[1] A. J. Tatem,[1,2,3] J. M. Cohen,[4] P. W. Gething,[5] G. Johnston,[6,7] R. Gosling,[8] R. Laxminarayan,[9] S. I. Hay,[3,5] D. L. Smith[3,6,7,9*]

When the Global Malaria Eradication Programme (GMEP) was launched in 1955 (*1*, *2*), all malaria-endemic countries outside of Africa were (or would soon be) eliminating malaria (*3*). The GMEP's design was based on a theory of malaria transmission dynamics and control that has become the standard for malaria elimination decisions today (*4*–*6*). When financial support for the GMEP collapsed in 1969, participating countries were caught at different stages of progress toward elimination (*1*). Examining their fate in the decades that followed provides a natural experiment that tests the theory. With a rise in funding (*7*) and renewed interest in eradication (*8*, *9*), there is now a need to revisit the lessons learned from the GMEP. We identify changes in the epidemiology of malaria when elimination is reached that could explain its stability and discuss how this calls for a reassessment of strategies for eradication.

Standard Theory of Malaria Elimination

Malaria elimination involves stopping transmission in a defined region until no parasites remain (*6*, *10*), through implementation of vector control, treatment of infected individuals, and other available interventions. Eradication is elimination at a global scale. Between 1945 and 2010, 79 countries eliminated malaria and 75 (95%) remained malaria-free (*11*), which shrank the geographical range of malaria (*12*). A recent review identified 75 resurgent malaria transmission events in 61 countries (*13*), which included marked increases in malaria incidence in 36 of the 49 countries (73%) that had participated in the GMEP but had failed to reach elimination (*11*, *14*). Resurgent malaria after the GMEP confirms some predictions of the standard theory, but elsewhere, the apparent stability of elimination is counterintuitive and may require a reappraisal of the theory. It suggests that elimination could be a way to solve the Sisyphean problem of malaria (*15*).

Standard theory for malaria elimination is based on a threshold concept called R_0 (*4*, *5*), the expected number of secondary infections produced by each infected human in the absence of control or acquired immunity. R_0 is a threshold condition because each case must cause at least one other case to have sustained transmission and endemic malaria. Mosquito aspects of R_0 are described by vectorial capacity, a single number that summarizes mosquito population density, longevity, blood-feeding habits, human-feeding preferences, and parasite development rates (*16*). Human aspects of R_0 are dominated by the long infectious period that is sustained by untreated asymptomatic infections lasting 6 to 8 months on average and, sometimes, longer than a year (*17*). Vector control reduces vectorial capacity; routine use of antimalarial drugs shortens infections; and, in the presence of such interventions, controlled reproductive numbers, denoted R_C, describe the reduced number of malaria cases arising from each malaria case (*9*). Elimination requires reducing R_C to <1, such that local transmission will eventually cease.



R_C estimates by year (1998–2002) for 27 GMEP countries that eliminated malaria. Elimination requires reducing R_C to <1. Full data spanning 1980–2010, with countries spelled out are available in SM. Mean R_C for available years spanning 1980–2010 is shown to the left.

After elimination, malaria importation poses a constant threat, because humans and mosquitoes carry malaria from endemic areas across international boundaries and within countries (*18*, *19*). In such settings, R_C describes probabilistically the total number of cases in an outbreak arising from an imported malaria case, so it is useful in planning the end of elimination (*10*) and the management of imported malaria.

According to standard malaria transmission theory, in order to eliminate malaria, R_C must be reduced to below 1, and as long as importation continues, control measures must be maintained such that R_C remains below 1 to prevent importation from reestablishing endemic malaria. Countries thus must retain the capability to stop imported malaria cases from transmitting, either by sustaining vector-control measures to keep R_C low even after elimination, or through intense surveillance to identify and cure all imported and subsequent infections before they can lead to resumption of transmission (*20*, *21*). If control measures are not sustained at sufficiently high levels, imported malaria could restart endemic transmission.

Malaria resurgence, which was common in those countries that tried, but failed, to eliminate malaria (*13*, *14*), conforms to this theory. The causes of resurgence were poorly documented but, most frequently, resurgence was blamed on the failure to sustain high intervention coverage levels (*13*). If this pattern held everywhere, elimination would require long-

[1]Emerging Pathogens Institute, University of Florida, Gainesville, FL 32610, USA. [2]Department of Geography and Environment, University of Southampton, Highfield, Southampton SO17 1BJ, UK. [3]Fogarty International Center, National Institutes of Health, Bethesda, MD 20892, USA. [4]Clinton Health Access Initiative, Boston, MA 02127, USA. [5]Spatial Ecology and Epidemiology Group, Department of Zoology, Oxford University, Oxford OX1 3PS, UK. [6]Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD 21205, USA. [7]Malaria Research Institute, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD 21205, USA. [8]Malaria Elimination Initiative, Global Health Group, University of California San Francisco, San Francisco, CA 94105, USA. [9]Center for Disease Dynamics, Economics and Policy, Washington, DC 20036, USA.

*Author for correspondence. E-mail: dlsmith@jhsph.edu

term investments in vector control until eradication is achieved. The best hope for eradication would be a massive campaign with sufficient funding to bring about a coordinated end of malaria transmission such that importation is no longer possible and control efforts can end. If true, eradication would require significant worldwide collaboration.

What Explains the Stability of Elimination?

Elimination countries once had endemic malaria (i.e., $R_0 > 1$), and they have remained malaria-free even though most of these countries have abandoned the elimination era vector-control programs. Without those interventions in place, why has transmission not resumed? Considering that all of these countries face ongoing malaria importation, the stability of elimination requires a closer look.

To assess contemporary malaria importation and to quantify R_C in elimination countries, national malaria reporting data were sought from 1980 onward for all countries that participated in the GMEP and eliminated malaria [see supplementary materials (SM)]. In many cases, the data were either sporadically or not collected or not publicly reported. Data were found for 30 countries that eliminated malaria, documenting 249,250 imported malaria cases over 596 country-years versus 4993 introduced (acquired by mosquito transmission from an imported case) or otherwise locally acquired. Formulas from branching theory were used to estimate R_C (*10*). The overall yearly average was $R_C \approx 0.04$, and ~85% (506 out of 596) of year-by-country R_C estimates were less than 0.01 (see the chart). Because these countries once sustained endemic transmission, this analysis suggests that, for 85% of country-years, transmission is proportionally lower by a factor of more than 100 (i.e., $R_0/R_C > 100$). Elimination has become highly stable.

This familiar, but often overlooked, pattern suggests that elimination has substantial advantages over the alternative strategy of minimizing the burden indefinitely without attempting to eliminate malaria, which faces the challenge of financing and distributing interventions to suppress transmission and to manage an ever-present but inapparent threat (*15*). Understanding the causes of very low R_C values in elimination countries is crucial to an assessment of the potential for malaria elimination to become a stable endpoint elsewhere. Important questions remain as to how the 99 countries with endemic malaria should weigh their technical and operational challenges (*20*, *21*) and assess the feasibility of and formulate plans for malaria elimination (*22*).

There are six main nonexclusive hypotheses for the stability of elimination after top-down control was ceased. (i) R_0 might have decreased because of indirect effects of economic development wholly unrelated to disease control, such as changes in mosquito ecology or human demography resulting in reduced human-vector contact; (ii) a decline in R_C might have caused economic development, which then reduced R_0 (*23*); (iii) R_C is low because countries have maintained vector-control programs (this is likely true only in few countries); (iv) R_C is low because country health systems have exerted highly effective control through routine surveillance and case management combined with outbreak control; (v) vector control permanently altered vector ecology and R_0; and (vi) the stability of elimination is partly due to travel patterns and the destinations of imported malaria. All six hypotheses could explain some portion of the 100-fold difference in R_C to varying degrees in different countries.

The appropriate plan for countries contemplating elimination depends on the causes of the stability. A critical determination is whether elimination was caused by external changes in R_0 unrelated to control efforts or whether elimination itself caused the permanent declines in R_0 and/or reductions in R_C. If malaria elimination requires an external change in R_0 (such as exogenous economic development), then control measures will need to be maintained until a transition occurs. If, instead, elimination itself causes stability through economic development and/or other means, then elimination is a highly desirable endpoint and should be aggressively pursued and supported. It is possible that one part of the mechanism by which development produces stability is through improved health systems and surveillance; if so, then improving health systems may be a necessary part of elimination.

Our analysis from post-elimination countries documenting large declines in R_0 or R_C—whatever the cause—suggests that eradication of malaria differs from smallpox and other vaccine-preventable, directly transmitted diseases (*24*). It has been argued that eradication is all-or-nothing, that interventions must remain in place until eradication has been achieved, and that there is no such thing as partial success (*6*). These patterns show that malaria elimination has been a partial success.

The possibility that the complete absence of ongoing malaria transmission can become a highly stable state is relevant for policy because it suggests that, before achieving global eradication, some countries could eliminate, scale back control measures, and rely on their health systems. Projected economic costs of elimination are dominated by management of imported malaria (*22*, *25*), but if elimination is stable, then it could save costs before achieving eradication. Objections to elimination have been drawn, ironically, from the experiences of countries that did not eliminate. This evidence suggests that elimination on path to eradication may be less risky than currently supposed, and that elimination could have substantial advantages over indefinite control.

If malaria elimination helps cause its own stability, then eradication may benefit from regional coordination, but it does not require a globally coordinated campaign. Malaria elimination can proceed like a ratchet, country-by-country and region-by-region culminating in global eradication.

References and Notes

1. G. Gramiccia, P. F. Beales, in *Principles and Practice of Malaria*, W. Wernsdorfer, I. McGregor, Eds. (Churchill Livingstone, Edinburgh, 1988), vol. 2, pp. 1335–1378.
2. J. A. Nájera *et al.*, *PLoS Med.* **8**, e1000412 (2011).
3. P. Yekutiel, in *Eradication of Infectious Diseases: A Critical Study*, M. A. Klingberg, Ed. (Karger, Basel, Switzerland, 1980), pp. 34–88.
4. G. MacDonald, *Bull. World Health Organ.* **15**, 613 (1956).
5. G. MacDonald, *Bull. World Health Organ.* **15**, 369 (1956).
6. World Health Organization, *Expert Committee on Malaria, Sixth Report* (WHO, Geneva, Switzerland, 1957).
7. D. M. Pigott, *et al.*, *Malar. J.* **11**, 246 (2012).
8. R. Feachem, O. Sabot, *Lancet* **371**, 1633 (2008).
9. Roll Back Malaria Partnership, *The Global Malaria Action Plan for a Malaria-Free World* (RBMP, WHO, Geneva, Switzerland, 2008).
10. J. M. Cohen *et al.*, *Malar. J.* **9**, 213 (2010).
11. UCSF Global Health Group in partnership with the Malaria Atlas Project, *Atlas of Malaria Eliminating Countries* (The Malaria Elimination Group, University of California, San Francisco Global Health Group, San Francisco, 2012).
12. S. I. Hay, C. A. Guerra, A. J. Tatem, A. M. Noor, R. W. Snow, *Lancet Infect. Dis.* **4**, 327 (2004).
13. J. M. Cohen *et al.*, *Malar. J.* **11**, 122 (2012).
14. R. G. Feachem *et al.*, *Lancet* **376**, 1517 (2010).
15. D. L. Smith *et al.*, *Science* **332**, 1384 (2011).
16. C. Garrett-Jones, *Bull. World Health Organ.* **30**, 241 (1964).
17. M. T. Bretscher *et al.*, *Epidemics* **3**, 109 (2011).
18. A. Wesolowski *et al.*, *Science* **338**, 267 (2012).
19. A. J. Tatem, D. L. Smith, *Proc. Natl. Acad. Sci. U.S.A.* **107**, 12222 (2010).
20. B. Moonen *et al.*, *Lancet* **376**, 1592 (2010).
21. B. Moonen *et al.*, *Malar. J.* **9**, 322 (2010).
22. Zanzibar Malaria Control Programme, *Malaria Elimination in Zanzibar: A Feasibility Assessment* (ZMCP, Zanzibar, Tanzania, 2009).
23. J. Sachs, P. Malaney, *Nature* **415**, 680 (2002).
24. J. G. Breman *et al.*, *PLoS Med.* **8**, e1000405 (2011).
25. O. Sabot *et al.*, *Lancet* **376**, 1604 (2010).

Acknowledgments: Authors received funding from the Bloomberg Family Foundation (D.L.S., G.J.), National Institute of Allergy and Infectious Diseases, NIH (U19AI089674; D.L.S., A.J.T., G.J.), the Bill & Melinda Gates Foundation (49446: D.L.S., A.J.T.; 1032350: A.J.T.; 1013170: J.M.C., R.G.), the RAPIDD program of the Science and Technology Directorate, Department of Homeland Security, and the Fogarty International Center, NIH (D.L.S., A.J.T., S.I.H.), and a Senior Research Fellowship from the Wellcome Trust (095066; S.I.H.). Authors also acknowledge support from the Malaria Elimination Group.

Supplementary Materials

www.sciencemag.org/cgi/content/full/339/6122/909/DC1

10.1126/science.1229509

ATMOSPHERIC SCIENCE

The Aerosol Nucleation Puzzle

Field studies in a boreal forest and laboratory experiments reveal how atmospheric aerosols are formed from gas molecules.

Meinrat O. Andreae

Even if the air around us seems perfectly clear, there are hundreds or thousands of tiny particles in every cubic centimeter of it. Where do they come from? Some, like dust and sea spray, are lofted from Earth's surface, but most aerosol particles are formed by condensation from molecules present in the gas phase. In a process called nucleation, compounds of low volatility, such as sulfuric acid, condense to small particles when they have reached a high enough concentration in the atmosphere, much like water condenses into cloud drops (see the figure, panel A). How this actually happens at the molecular level has long been an enigma of atmospheric science. Like the shells of a Russian doll, scientists have peeled away layer after layer of this mystery by developing techniques to observe smaller and smaller particles (*1*). On page 943 of this issue, Kulmala *et al.* (*2*) report on their observations of aerosol nucleation in a remote forest in Finland (see the figure, panel B), using new analytical tools to reveal how gas molecules join together to form particles.

Experimental observation of these processes has been hindered because nucleation occurs in an instrumental blind spot between the realm of molecules and that of particles. For example, atmospheric pressure inlet mass spectrometers (API-MS) detect molecules up to several hundred Daltons, corresponding to particles up to about 1-nm diameter, whereas particle detectors typically are limited to sizes of a few nanometers or larger. New instruments have closed this gap: Chemical ionization API time-of-flight MS can now detect atmospheric molecules and particles reaching sizes up to a few nanometers, whereas particle size magnifiers and neutral cluster and air ion spectrometers reach down to sizes ~1 nm. The chemical composition of somewhat larger particles (4 to 7 nm) can be determined with thermal desorption chemical ionization MS. Deployment of these and related techniques, together with a host of supporting measurements, have now produced a consistent picture of the events that lead from molecules to particles.

Biogeochemistry Department, Max Planck Institute for Chemistry, Mainz, D-55020 Germany. E-mail: m.andreae@mpic.de

From gases into particles. (**A**) Atmospheric nucleation is represented for sulfuric acid (sulfur atoms in yellow) and organic molecules (carbon atoms in green) joining to form a critical cluster, which grows to become an aerosol particle (*8*). (**B**) The boreal forest at Hyytiälä, Finland, where Kulmala *et al.* performed detailed field measurements of aerosol formation. (**C**) The CLOUD experimental chamber at CERN, Geneva, provides complementary laboratory studies of aerosol formation.

Measurements at Hyytiälä, a remote site in the boreal forest of Finland, suggest that nucleation proceeds in three stages. The first is characterized by small clusters, up to about 1.2-nm diameter, that are present all the time at concentrations of a few thousand per cubic centimeter air, and that are not strictly part of the nucleation process. These clusters appear to represent a state close to dynamic equilibrium, where molecules rapidly condense and evaporate, resulting in very slow net growth. There is widespread consensus that sulfuric acid is the critical species at this stage, and enough sulfuric acid (H_2SO_4) is present in the ambient air at Hyytiälä to account for the formation of these clusters. Growth accelerates in the second stage (1.2 to 1.7 nm), but quantum-chemical calculations indicate that species other than H_2SO_4 must participate to stabilize these growing clusters. Recent work suggests that ammonia, organic amines, and oxidized organic molecules from biogenic hydrocarbons help bind the H_2SO_4 molecules together.

At this stage, the nascent particles undergo one of two fates. On about half of the days investigated at Hyytiälä, formation and growth of clusters was compensated by their loss to the surface of particles already present in the air, resulting in only very small numbers of particles growing to sizes larger than 2 nm. In contrast, on so-called "event days," cascading growth of clusters larger than about 1.5 nm led to nucleation bursts, with thousands of particles >2 nm being present in a cubic centimeter of air during the late morning, the photochemically most active time of day.

From the physicist's point of view, these are exciting results, because they provide a glimpse of nucleation in action by peering into the critical size range below about 2 nm. From the chemist's perspective, however, mysteries still remain. What actually is the role of H_2SO_4, H_2O, and the various potential "helper" molecules, such as amines and organics, in the cluster regime between about 0.5 nm, the size of a single H_2SO_4 molecule, and the range above ~1.5 nm, where organic

molecules dominate growth, at least at this remote forest site? Laboratory experiments may be able to provide some answers. In the most sophisticated laboratory study of the nucleation process so far, researchers brought together H_2SO_4 and organic molecules in the CLOUD (Cosmics Leaving Outdoor Droplets) chamber at CERN at environmentally relevant concentrations (see the figure, panel C). They followed, step-by-step, how molecular species come together in orderly progression and grow into clusters (*3*). The interplay of field and laboratory studies must continue in order to improve our understanding of the interface between molecules and particles.

What makes "event" days different from "nonevent" days? Is it just the competition between the rate at which H_2SO_4 molecules are produced and the speed at which the clusters and nucleating particles are consumed by preexisting particles before they can grow on their own? Or does the abundance and composition of the "helper" molecules play a critical role? Amines may play a role in binding the clusters together, at least in some environments (*4*). "Organics" are important, especially in the growth stage, but precisely what organics? Here is where the enigma of nucleation connects with the mystery of the chemical processes that lead from volatile organic gases to organic aerosols in the atmosphere.

Finally, how representative is a forest in the middle of Finland for the global atmosphere? We know, for example, that in the Amazon forest, nucleation almost never occurs (*5*). Solving the nucleation puzzle still requires pieces from a large variety of natural and human-influenced environments before we can understand the whole picture. This understanding is needed to assess the role that nucleation and new particle formation play in the budget of the atmospheric aerosol (*6*) and its impact on climate through its effects on cloud formation and the radiation budget (*7*).

References

1. R. Y. Zhang, A. Khalizov, L. Wang, M. Hu, W. Xu, *Chem. Rev.* **112**, 1957 (2012).
2. M. Kulmala *et al.*, *Science* **339**, 943 (2013).
3. H. Keskinen *et al.*, *Atmos. Chem. Phys. Discuss.* **12**, 31071 (2012); www.atmos-chem-phys-discuss.net/12/31071/2012
4. J. N. Smith *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **107**, 6634 (2010).
5. C. Pöhlker *et al.*, *Science* **337**, 1075 (2012).
6. J. Merikanto, D. V. Spracklen, G. W. Mann, S. J. Pickering, K. S. Carslaw, *Atmos. Chem. Phys.* **9**, 8601 (2009).
7. J. Kazil *et al.*, *Atmos. Chem. Phys.* **10**, 10733 (2010).
8. R. Zhang, *Science* **328**, 1366 (2010).

10.1126/science.1233798

IMMUNOLOGY

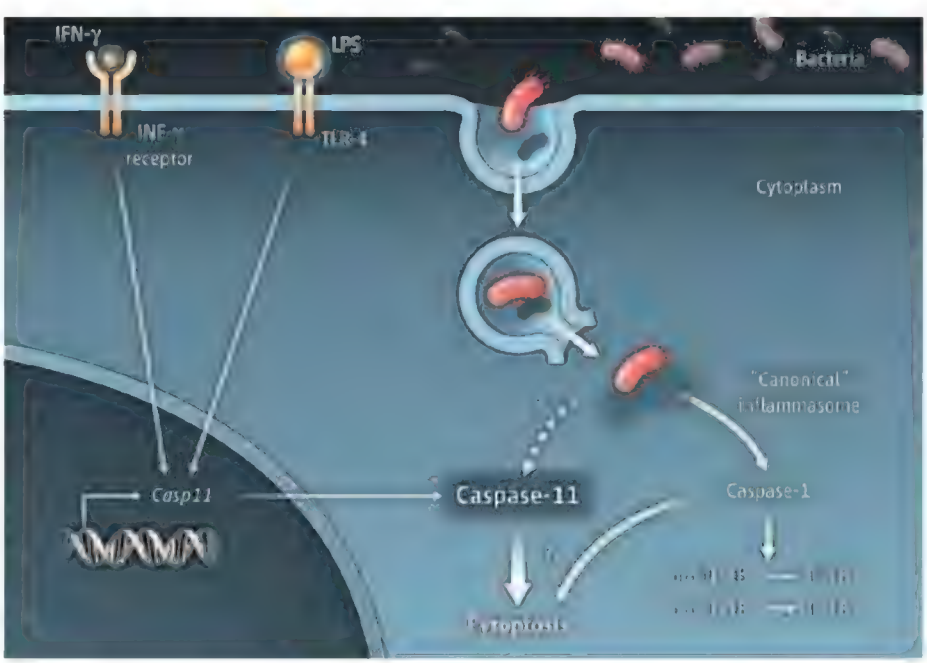
Bacterial Escape Artists Set Afire

An enzyme protects mice from bacterial infection by inducing pyroptosis, an inflammatory form of cell death.

Marija Cemma[1,2] and John H. Brumell[1,2,3]

Bacteria can occupy multiple compartments within host cells during infection, including vacuoles (*1*). Some bacteria manage to escape from vacuoles and enter the cytosol. These so-called "cytosol-adapted" pathogens can sustain rapid growth in the nutrient-rich environment. Host cellular defenses that protect the cytosol from invading bacteria include phagosome repair mechanisms (*2*) and autophagy (*1*). On page 975 of this issue, Aachoui *et al.* (*3*) report that the enzyme caspase-11 can also protect mice from a lethal infection by promoting a type of programmed cell death called pyroptosis, a response that may be a last line of defense against pathogens that escape into the cytosol.

Caspases are cysteine proteases that have many roles, including the promotion of inflammation (*4*). The role of caspase-1 in this context is well established. It is activated by a multiprotein complex called an inflammasome that detects bacterial products in the cytosol. Once activated, caspase-1 processes proinflammatory cytokine precursors, prointerleukin1-β (pro–IL-1β) and pro–IL-18, into their active forms (*4*). Caspase-1 also triggers pyroptosis, a form of cell death characterized by cell swelling, loss of plasma membrane integrity, and cell lysis, with subsequent release of proinflammatory cytokines (*5*) (pyroptosis, translated from Greek, is "the falling of fire," referring to the release of proinflammatory factors). It was recently shown that caspase-11 triggers caspase-



Protection from pyroptosis. During infection, expression of caspase-11 is up-regulated through Toll-like receptor 4 and/or IFN-γ receptor signaling. After bacterial escape from vacuoles, caspase-11 promotes pyroptosis, leading to bacterial clearance and host protection. It is unclear how caspase-11 detects cytosolic bacteria (dashed line), but a "noncanonical inflammasome" has been proposed. Caspase-1 activation by canonical inflammasomes also promotes pyroptosis under some conditions, in addition to mediating pro-inflammatory cytokine release.

CREDIT: C. BICKEL/*SCIENCE*

[1]Cell Biology Program, Hospital for Sick Children, Toronto, Ontario M5G 1X8, Canada. University of Toronto, Toronto, Ontario M5S1A8, Canada. [2]Department of Molecular Genetics, and [3]Institute of Medical Science, University of Toronto, Toronto, Ontario, M5S1A8, Canada. *E-mail: john.brumell@sickkids.ca

1–independent pyroptosis, as well as caspase-1–dependent cytokine production, in response to Gram-negative bacteria (*6*–*10*).

Aachoui *et al.* establish that caspase-11 has a protective role during infection with cytosolic bacteria. The authors show that mice lacking both caspase-1 and -11 (*Casp1*$^{-/-}$*Casp11*$^{-/-}$) succumb to infection by the cytosol-adapted Gram-negative pathogen *Burkholderia thailandensis*. By contrast, mice lacking functional inflammasomes (through the deletion of genes encoding the inflammasome components NLRP3, NLRC4, and ASC), and that also have impaired IL-1β secretion, were resistant to *B. thailandensis*, indicating that resistance does not depend on cytokine secretion. The *Casp1*$^{-/-}$*Casp11*$^{-/-}$ macrophages did not activate cell death by pyroptosis in response to the bacterial infection, unlike macrophages lacking inflammasomes. Aachoui *et al.* demonstrate that caspase-11 is primarily responsible for pyroptosis and that *Casp11*$^{-/-}$ mice recapitulate the lethal phenotype of *Casp1*$^{-/-}$ *Casp11*$^{-/-}$ mice with *B. thailandensis* infection. Caspase-11–dependent pyroptosis resembled caspase-1–dependent pyroptosis, as judged by membrane damage as well as glycine sensitivity (glycine blocks ion fluxes in damaged eukaryotic cells, thereby preventing swelling and lysis in pyroptosis triggered by caspase-1).

Aachoui *et al.* further support their finding that caspase-11 protects against cytosolic bacteria by using the vacuole-adapted pathogen, *Salmonella typhimurium*. A mutant strain, Δ*sifA*, localizes to the cytosol because it does not secrete SifA, a "type 3 secreted effector" protein that maintains the integrity of *Salmonella*-containing vacuole (*11*). Only in the presence of caspase-11 can the wild-type bacteria substantially outcompete the Δ*sifA* mutant. This implies that wild-type *S. typhimurium* residing in the vacuole that are not exposed to caspase-11 are protected, whereas the cytosol-localized Δ*sifA* mutant is vulnerable to caspase-11 detection and subsequent elimination by pyroptosis.

It is noteworthy that the authors observe induction of caspase-11 expression after priming by bacterial lipopolysaccharide (LPS) or interferon-γ (IFN-γ) a cytokine that is critical for immunity against intracellular bacterial infections. Caspase-11 may be constitutively active upon its expression (*10*). However, the finding that bacterial entry to the cytosol is required for induction of pyroptosis suggests that other signals are required for caspase-11 to execute its function. Whether caspase-11 undergoes post-translational modifications or subcellular targeting following bacterial escape into the cytosol is not yet known.

Caspase-11 is now recognized as an important part of the inflammatory arsenal that can be either protective (*3*) or damaging depending on the physiological context of its activity (*6*, *7*). Yet, the study of Aachoui *et al.* raises two major questions. What is the signal(s) that activate caspase-11? A "noncanonical inflammasome" has been proposed to activate caspase-11 (*6*), but no molecular players have yet been identified. And what are the substrates of caspase-11 that promote pyroptosis? These gaps in understanding of caspase-11 function will be important topics for future studies.

There is a dynamic interplay between caspases and bacterial pathogens. Some bacteria use caspases to promote their pathogenesis. For example, *Listeria monocytogenes* exploits caspase-7 to repair bacteria-induced damage to the plasma membrane to preserve its replicative niche (*12*). Similarly, *S. typhimurium* exploits caspase-3 to cleave and activate a virulence factor, SipA (*13*). By contrast, other pathogens inhibit inflammosome activation to attenuate the immune response. Bacterial factors from *Mycobacterium tuberculosis* and *Yersinia pestis* promote their virulence by inhibiting inflammasome activation, thereby dampening their host's immune response (*14*, *15*). In the case of *Y. pestis*, its type 3 secreted effector, YopM, sequesters caspase-1 into the nucleus where it can no longer interact with the inflammasome (*14*). It remains to be elucidated whether some cytosol-adapted pathogens, like *L. monocytogenes* and *Shigella flexneri*, have devised strategies to evade or exploit casapase-11–dependent pyroptosis.

References

1. M. Cemma, J. H. Brumell, *Curr. Biol.* **22**, R540 (2012).
2. M. J. Davis, B. Gregorka, J. E. Gestwicki, J. A. Swanson, *J. Immunol.* **189**, 4488 (2012).
3. Y. Aachoui *et al.*, *Science* **339**, 975 (2013); 10.1126/science.1230751.
4. A. S. Yazdi, G. Guarda, M. C. D'Ombrain, S. K. Drexler, *J. Innate Immun.* **2**, 228 (2010).
5. E. A. Miao, J. V. Rajan, A. Aderem, *Immunol. Rev.* **243**, 206 (2011).
6. N. Kayagaki *et al.*, *Nature* **479**, 117 (2011). C.
7. P. Broz *et al.*, *Nature* **490**, 288 (2012).
8. C. L. Case *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **110**, 1851 (2013).
9. P. Gurung *et al.*, *J. Biol. Chem.* **287**, 34474 (2012).
10. V. A. Rathinam *et al.*, *Cell* **150**, 606 (2012).
11. C. R. Beuzón *et al.*, *EMBO J.* **19**, 3235 (2000).
12. S. K. Cassidy *et al.*, *PLoS Pathog.* **8**, e1002628 (2012).
13. C. V. Srikanth *et al.*, *Science* **330**, 390 (2010).
14. C. N. LaRock, B. T. Cookson, *Cell Host Microbe* **12**, 799 (2012).
15. S. S. Master *et al.*, *Cell Host Microbe* **3**, 224 (2008).

10.1126/science.1235639

BIOCHEMISTRY

Integrative Structural Biology

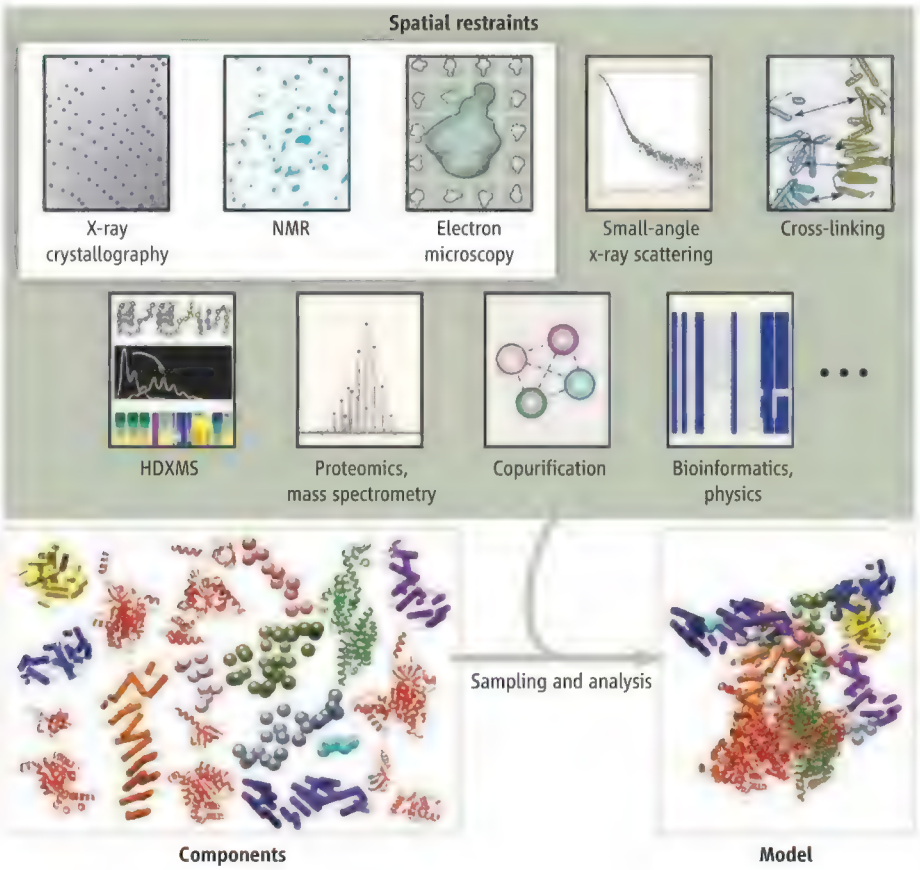
Andrew B. Ward,[1] Andrej Sali,[2] Ian A. Wilson[1,3]

Integrative approaches using data from a wide variety of methods are yielding model structures of complex biological assemblies.

Biological assemblies and machines often elude structural characterization, hampering our understanding of how they function, how they evolved, and how they can be modulated. A number of macromolecular assemblies have been reconstructed over the years by piecemeal efforts, such as fitting high-resolution crystal structures of individual components into lower-resolution electron microscopy (EM) reconstructions of the entire complex (*1*). Although notable successes have been achieved in this way, ambiguous or conflicting models can still arise (*2*–*4*). Thus, structural and computational biologists have been looking for new ways to put all of the pieces back together. Sophisticated integrative approaches are being developed (*5*, *6*) that combine information from different types of experiments, physical theories, and statistical analyses to compute structural models of multicomponent assemblies and complex biological systems.

In addition to the conventional biophysical techniques of x-ray crystallography, nuclear magnetic resonance (NMR) spectroscopy, EM, and small-angle x-ray scattering (SAXS), a growing number of experimental methods can also provide valuable infor-

[1]Department of Integrative Structural and Computational Biology, International AIDS Vaccine Initiative (IAVI) Neutralizing Antibody Center, and Center for HIV/AIDS Vaccine Immunology and Immunogen Discovery (CHAVI-ID), Scripps Research Institute, La Jolla, CA 92037, USA. [2]Department of Bioengineering and Therapeutic Sciences, Department of Pharmaceutical Chemistry, and California Institute for Quantitative Biosciences (QB3), University of California, San Francisco, CA 94158, USA. [3]Skaggs Institute for Chemical Biology, Scripps Research Institute, La Jolla, CA 92037, USA. E-mail: abward@scripps.edu



Complex structure solutions. Models of macromolecules and their complexes can be constructed by combining different types of information generated by various experimental and theoretical techniques (gray box). The data are converted into spatial restraints, which are combined into a scoring function that guides sampling algorithms to obtain a detailed structural model.

mation about the structures and dynamics of proteins and their assemblies. These methods include sequence comparisons of related proteins, copurification, hydrogen-deuterium exchange mass spectrometry (HDXMS), single-molecule fluorescence, atomic force microscopy, analytical spectroscopy (both electron paramagnetic resonance and double electron-electron resonance), light scattering, chemical cross-linking, and mutagenesis (see the figure).

The individual pieces of data gathered using different techniques can provide invaluable restraints on the conformation, position, and orientation of the components in an assembly or biological system (*5*). Relative to the use of any single set or type of data, simultaneous use of all such restraints can markedly improve the accuracy, precision, and completeness of a model, especially when high-resolution structural data on the entire complex are not available.

Because of the many degrees of freedom in macromolecular structures and the difficulty of combining disparate data, models must be computed with algorithms that sample as many potential solutions as possible given the computing power available. These algorithms are driven by a scoring function consisting of the individual spatial restraints and are analogous to methods used in x-ray crystallography and NMR spectroscopy, which also generate models by minimizing differences between experimental data and data calculated from a model. Assessing how to best combine and weigh different types of data from multiple sources is a prerequisite for constructing structural models of increasingly larger and more dynamic macromolecular complexes.

A useful test of a model is whether it explains all data points within their own error bars and whether the entire data set is redundant, meaning that a subset of the data can be omitted without any significant impact on the model. In such a case, the confidence in the model, the data, and the parameters used for modeling can be high. When a subset of the data points cannot be satisfied by a single model because the data were collected from a heterogeneous sample and/or the data are noisy, more sophisticated methods for combining individual restraints are needed. In such cases, emphasis is placed on evaluating models in an objective manner, using Bayesian (*7*) and other statistical methods that explicitly take into account the noise in the data and/or multiple structural states in the sample.

Integrative, restraint-based approaches can be used whenever a challenging structural biology problem is encountered, from an individual protein to a small macromolecular machine to a large multicomponent cellular assembly. Thus, integrative approaches span wide resolution ranges and bridge observations made from the atomic to the cellular level. The following three examples illustrate the power of these new methods in generating models at different levels of resolution.

Some of the most successful applications of integrative approaches have resulted from combining sparse experimental observations with computation to generate atomic-level models of macromolecules. Rosetta (*8*), a platform for modeling protein structures, works by exhaustive calculations under a set of assumptions about the underlying geometry and chemistry of peptides. These assumptions reduce the nearly infinite sampling necessary to fold a one-dimensional sequence of amino acids into a three-dimensional shape. Experimental restraints from NMR (*9*) or EM (*10*) can further narrow the search and help to converge on more accurate models. For example, Loquet *et al.* used solid-state NMR, EM, and Rosetta to build an atomic-level model of the bacterial type III secretion needle used to inject its proteins into host cells (*11*). The model revealed details of the supramolecular interfaces of the component protomers, providing a structural understanding of this machine that had eluded characterization by single techniques.

Two recent independent studies of the molecular architecture of the 26*S* proteasome exemplify the value of integrative approaches for medium-resolution structures. Lander *et al.* combined EM reconstructions and x-ray crystal structures (*12*), whereas Lasker *et al.* used restraints from a variety of data sets (EM, x-ray crystallography, chemical cross-linking, and proteomics) and employed the Integrative Modeling Platform package (*6*, *13*) to build an almost identical model of the 26*S* proteasome (*14*). Lasker *et al.*'s model was further tested by systematically removing some input data, recalculating a model, and assessing it against the omitted data. Although neither model resolved all interactions at an atomic level, they provided a detailed understanding of the arrangement of the component subunits and were therefore

extremely informative about the evolution and function of the 26*S* proteasome.

At low resolution, chromatin has also been modeled through integrative approaches. In this way, Duan *et al.* constructed a three-dimensional model of the yeast genome (*15*), uncovering the topology and spatial relationships of different chromosomal elements. For this study, the restraints were garnered from cross-linking, restriction enzyme digestion, ligation, and deep sequencing, thereby revealing the three-dimensional structure of the genome at a level of detail not accessible to any conventional imaging method typically used to study assemblies of this size. These inferential, cellular-scale approaches enable comparison of normal and aberrant cells and may eventually serve an important diagnostic role in medicine.

As integrative methods evolve, structural models can then be revisited and new data incorporated, so that the model can be continuously improved and revised using the latest information (*13*). Integrative software tools should therefore be flexible enough to incorporate new data and/or restraints. The Protein Data Bank (www.pdb.org) is facilitating this process by acting as a curator for a variety of structural data from different methods as well as models based on these data.

Any experimental observation can in principle be converted into a restraint for building ever more complex models. The reach and impact of structural biology can thus be extended to a wider and more diverse audience. Using these new computational and bioinformatics approaches to collect and integrate diverse pieces of structural and experimental data, Humpty Dumpty can be put back together again.

References and Notes

1. G. C. Lander, H. R. Saibil, E. Nogales, *Curr. Opin. Struct. Biol.* **22**, 627 (2012).
2. H. W. Wang *et al.*, *Nat. Struct. Mol. Biol.* **16**, 1148 (2009).
3. P. W. Lau *et al.*, *Nat. Struct. Mol. Biol.* **19**, 436 (2012).
4. R. Henderson *et al.*, *Structure* **20**, 205 (2012).
5. F. Alber, F. Förster, D. Korkin, M. Topf, A. Sali, *Annu. Rev. Biochem.* **77**, 443 (2008).
6. D. Russel *et al.*, *PLoS Biol.* **10**, e1001244 (2012).
7. W. Rieping, M. Habeck, M. Nilges, *Science* **309**, 303 (2005).
8. www.rosettacommons.org
9. N. G. Sgourakis *et al.*, *J. Am. Chem. Soc.* **133**, 6288 (2011).
10. F. DiMaio, M. D. Tyka, M. L. Baker, W. Chiu, D. Baker, *J. Mol. Biol.* **392**, 181 (2009).
11. A. Loquet *et al.*, *Nature* **486**, 276 (2012).
12. G. C. Lander *et al.*, *Nature* **482**, 186 (2012).
13. http://integrativemodeling.org
14. K. Lasker *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **109**, 1380 (2012).
15. Z. Duan *et al.*, *Nature* **465**, 363 (2010).

Acknowledgments: Supported by the Protein Structure Initiative of the National Institute of General Medical Sciences, the CHAVI-ID and HIVRAD (HIV Vaccine Research and Design) programs of the National Institute of Allergy and Infectious Diseases, and IAVI.

10.1126/science.1228565

PLANT SCIENCE

Preservation of Recalcitrant Seeds

Cryogenic technologies help to preserve plant biodiversity in seed banks, particularly in the tropics.

Christina Walters,[1] Patricia Berjak,[2] Norman Pammenter,[2] Kathryn Kennedy,[3] Peter Raven[4]

Concerns about the rapid erosion of plant diversity have spawned a host of seed-banking initiatives (*1*). These repositories provide critical germ plasm needed to understand, maintain, and manage natural variation within and among species (*2*). However, numerous plant species and much of the humid tropics are underserved in these endeavors because of the perceived problem of seed recalcitrance (*3*). About 75 to 80% of angiosperm species (*4*, *5*) produce orthodox seeds that can survive drying and prolonged storage at –20°C. By contrast, 5 to 10% of angiosperm species produce recalcitrant seeds that do not survive desiccation (*3*) and are killed in the freezer when ice crystals form. How can their preservation be ensured?

These seeds are called recalcitrant because they die rapidly when treated under the same genebanking conditions used to store orthodox seeds. Recalcitrant seeds, like most living organisms but unlike orthodox seeds, need water to survive. By nature, recalcitrant seeds are short-lived; they either germinate or are eaten by animals in the wild (*3*). "Intermediate" seeds, produced by 10 to 15% of angiosperm species (*4*), withstand sufficient dehydration to prevent formation of lethal ice; nevertheless, their seeds are short-lived in the freezer for unknown reasons (*6*).

The misperception that recalcitrant seeds cannot be stored arises from the assumption that there is only one storage strategy at our disposal—standard freezers—to manage the wide variation in storage physiology exhibited by plant seeds. In reality, long-term storage of recalcitrant and intermediate seeds (or seed parts) is possible with cryogenic technologies (*7*–*9*). Because cryogenic storage requires specialized infrastructure and personnel and is costly, most genebanks, even some of the newer ones, invest only in freezer storage.

Cryopreservation involves storage at ultralow temperature, often in liquid nitrogen (–196°C) (*7*–*9*). Rapidly advancing methods can be used to essentially stop water from freezing within recalcitrant seed cells and obviate lethal freezing damage (*3*, *7*–*9*). The technology requires carefully dehydrating tissues and cooling them extremely rapidly (*7*). Small samples are required to obtain diffusion and heat transfer rates needed to prevent ice formation (*7*).

The problem is that recalcitrant seeds tend to be large compared with orthodox or intermediate seeds (see the figure) (*5*). In fact, most recalcitrant seeds are far too large to be rapidly dehydrated or cooled effectively when exposed to liquid nitrogen. The major breakthrough that led to successful cryopreservation of recalcitrant germ plasm was the ability to surgically dissect out the growing portion of the seed (termed the embryonic axis) and germinate it in vitro (*10*). Recovery may be enhanced by exposing embryonic axes to cryoprotectants (substances that protect against dehydration and freezing damage) (*8*, *9*) and reducing stress-related free radical–mediated damage (*11*).

Cryogenic technology has been successfully refined for embryonic axes of a spectrum of temperate recalcitrant-seeded species (*8*). Intermediate seeds are smaller and drier and do not even require these exacting preparations. However, unmet challenges still limit successful cryopreservation of embryonic axes of some recalcitrant seeds from the tropics and subtropics (*11*). The seeds of these species have embryonic

[1]U.S. Department of Agriculture–Agricultural Research Service National Center for Genetic Resources Preservation, Fort Collins, CO 80521, USA. [2]Plant Germplasm Conservation Research, School of Life Sciences, University of KwaZulu-Natal, Durban, 4001 South Africa. [3]The Center for Plant Conservation, St. Louis, MO 63166–0299, USA. [4]Missouri Botanical Garden, St. Louis, MO 63166–0299, USA. E-mail: christina.walters@ars.usda.gov

Preservation challenge. Examples of recalcitrant (avocado, middle), intermediate (papaya, left), and orthodox seeds (melon, right). The avocado fruit has a single large seed with a visible embryonic axis (arrow) that continues to grow and germinate within the seed after the fruit is shed. Melon and papaya seeds can be separated from the fruit and dried to maintain viability in the freezer (melon) or in liquid nitrogen (papaya). Similar treatment is lethal to avocado.

axes that are metabolically active and poised for quick germination. Several species have large embryonic axes that place constraints on diffusion and energy exchange rates. Encouragingly, though, greater cryopreservation success for axes of these species is forthcoming through research to reduce excision damage, lower metabolic activity, and increase totipotency of small axis sections. This research should be prioritized to meet urgent conservation challenges.

In addition to misperceptions about the futility of genebanking recalcitrant seeds, delays in plant genebanking initiatives—often in areas with especially high levels of endemism and rarity—also arise from an assumption that most plants from warm, humid climates produce recalcitrant seeds. Globally, 20 to 25% of angiosperm and gymnosperm species produce nonorthodox seeds (*4*, *5*). However, in tropical evergreen rainforests, up to 47% of the indigenous flora is hypothesized to produce recalcitrant seeds (*5*, *12*, *13*). Confirming these estimates will require extensive surveys of the postharvest physiology of seeds from various regions. Such surveys for the Hawaiian Islands revealed that only ~3% of the endemic species actually produce recalcitrant seeds (*14*). Concerted efforts are now finally under way to bank seeds of this highly imperiled flora.

The condition of seed recalcitrance can confound additional challenges of conserving genetic diversity of wild plant species ex situ. For example, wild populations are often remote, and individuals within populations may be separated by large distances. Flowering, seed fill, and fecundity can be highly variable, making it difficult to obtain samples of the desired size or quality (*15*, *16*). Repeated sampling is therefore necessary to acquire genetically representative samples ex situ. Cryopreservation, then, becomes the vehicle to store germ plasm while adequate collections can be made to serve the purposes of conservation.

Many species producing recalcitrant seeds are trees (*5*), and genetic diversity of tree species is especially difficult to maintain ex situ. Regeneration of dwindling or aging germ plasm from trees requires more land and time as a result of their large size and extended juvenility period. Also, for many plant groups but especially for trees, preventing hybridization with closely related species may require controlled pollinations to maintain taxonomic identity and traits in offspring populations. Seed banking using cryopreservation extends longevity, thereby reducing the frequency of regeneration and risks of genetic erosion. Knowledge of the population structure (*16*) and additional cryopreserved collections of pollen and other tissues as germ plasm (*17*, *18*) can be invaluable tools to optimize space and maintain genetic diversity in tree species. Effective ex situ conservation for plants must make use of the full array of genebanking tools.

The cryogenic storage of nonorthodox germ plasm is considerably more costly than the more conventional freezer storage used for orthodox seeds. Nevertheless, cryopreservation is far less expensive and safer than the alternative approach of maintaining these genetic resources as growing plants in tissue culture, greenhouses, or field plantings (*19*, *20*). Also, once embryonic axes are cryobanked, maintenance costs are fairly low (*19*, *20*). Cryopreservation thus represents the most efficient method to preserve viability of genetically diverse populations (*8*).

References

1. E. Pennisi, *Science* **329**, 1274 (2010).
2. C. Walters, G. M. Volk, C. M. Richards, *Biodiversity* **9**, 68 (2008).
3. P. Berjak, N. W. Pammenter, *Ann. Bot. (Lond.)* **101**, 213 (2008).
4. J. B. Dickie, H. W. Pritchard, in *Desiccation and Plant Survival*, M. Black, H. W. Pritchard, Eds. (CABI, Wallingford, UK, 2002), pp. 239–260.
5. J. C. Tweddle, J. B. Dickie, C. C. Baskin, J. M. Baskin, *J. Ecol.* **91**, 294 (2003).
6. T. D. Hong, R. H. Ellis, "A protocol to determine seed storage behaviour." *IPGRI Technical Bulletin* No. 1 (International Board for Plant Genetic Resources, Rome, Italy, 1996).
7. C. Walters *et al.*, in *Plant Cryopreservation: A Practical Guide*, B. M. Reed, Ed. (Springer, New York, 2008), pp. 465–484.
8. F. Engelmann, *Methods Mol. Biol.* **710**, 155 (2011).
9. E. E. Benson, *Crit. Rev. Plant Sci.* **27**, 141 (2008).
10. M. N. Normah, H. F. Chin, Y. L. Hor, *Arg. Pertanika* **9**, 299 (1986).
11. P. Berjak, B. Sershen, B. Varghese, N. W. Pammenter, *Seed Sci. Res.* **21**, 187 (2011).
12. K. N. Hamilton, C. A. Offord, P. Cuneo, M. A. Deseo, *Plant Species Biol.* **28**, 51 (2013).
13. M. I. Daws, N. Garwood, H. W. Pritchard, *Ann. Bot. (Lond.)* **97**, 667 (2006).
14. L. Weisenberger, L. M. Hill, A. Yoshinaga, K. Wood, C. Walters, *Pol. J. Nat. Sci.* **5** (suppl.), 281 (2008).
15. A. T. Kramer, K. Havens, *Trends Plant Sci.* **14**, 599 (2009).
16. C. M. Richards *et al.*, *Tree Genet. Genomes* **5**, 339 (2009).
17. B. Reed, V. Sarasan, M. Kane, E. Bunn, V. Pence, *In Vitro Cell. Dev. Biol. Plant* **47**, 1 (2011).
18. L. E. Towill, in Ex Situ *Plant Conservation: Supporting Species Survival in the Wild*, E. O. Guerrant, K. Havens, M. Maunder, Eds. (Island Press, Washington, DC, 2004), pp. 180–188.
19. V. Pence, *In Vitro Cell. Dev. Biol. Plant* **47**, 176 (2011).
20. D. Z. Li, H. W. Pritchard, *Trends Plant Sci.* **14**, 614 (2009).

10.1126/science.1230935

APPLIED PHYSICS

Graphene Nanophotonics

Graphene allows us to control light in innovative ways, providing the prospect of unprecedented nanophotonic devices.

F. Javier García de Abajo

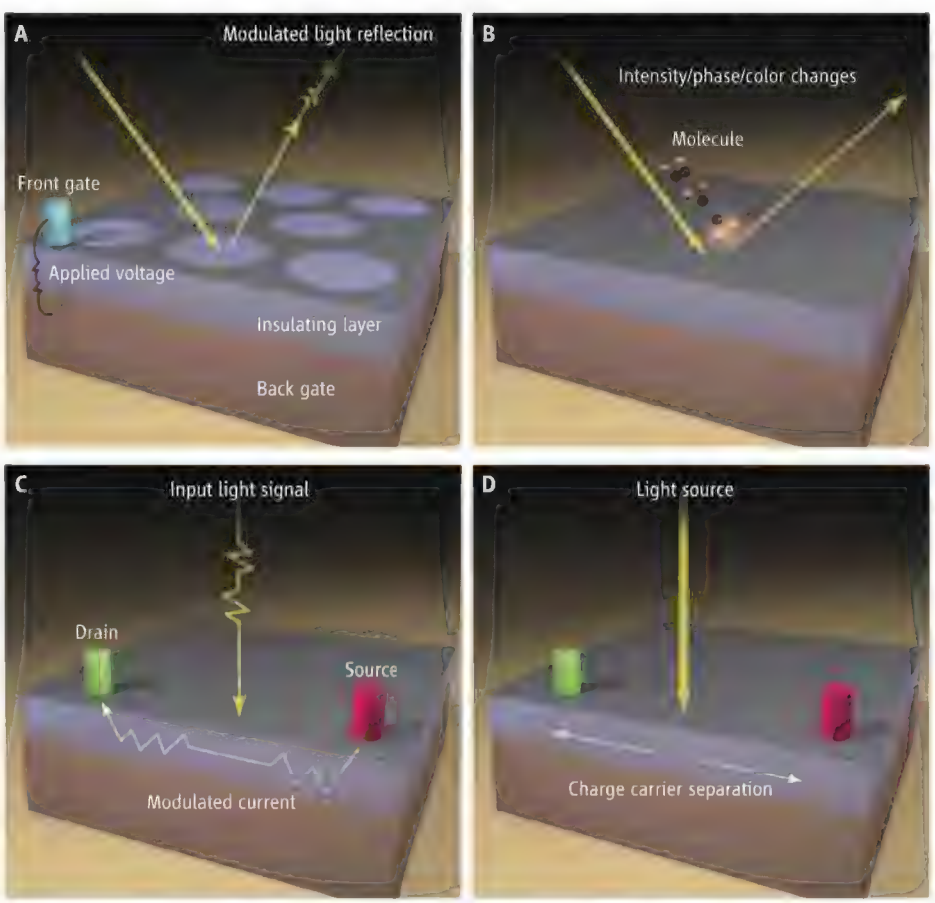
Nanophotonics—the control of light at nanometer scales—has produced some amazing feats during the past two decades, using light to trap and manipulate small objects such as living cells, to propagate and process information faster than conventional electronic computers, and to reveal the presence of minute amounts of hazardous substances and pathogenic agents. Advances in precise nanofabrication and self-assembly have enabled the routine production of engineered nanostructures that can confine light to regions much smaller than the optical wavelength, thus enhancing the light intensity by several orders of magnitude at designated locations. To complete the picture, fast ways of modulating the optical response of these nanostructures have been pursued, but with mixed success only. The availability of new materials with novel photonic properties is a welcome stimulus to this challenging task, with graphene—a one-atom-thick layer of carbon atoms—offering great prospects for meeting these challenges.

Although high-quality graphene has been available for less than a decade, it has already made an impressive debut in nanophotonics. In addition to its ability to confine light to small volumes, its optical response can be modulated by subjecting it to electrical potentials, thereby allowing the manipulation of light at fast microelectronics speeds. Graphene can be integrated into devices operating under ambient conditions, it is structurally robust, it is electrically and chemically tunable, it exhibits an unusually strong nonlinear response, it can act both as a highly transparent thin conductor and as a perfect light absorber (depending on how it is prepared), and it is widely available in high quality at moderate cost. For these reasons, graphene has raised high expectations for its exploitation in a vast range of nanophotonic devices (*1*) and for the development of a new paradigm of tunable photonic structures that can control and process optical signals (*2*). Funding at an unprecedented level is being committed to support and extend these efforts (*3*).

Instituto de Química Física Rocasolano, Consejo Superior de Investigaciones Científicas (IQFR-CSIC), Serrano 119, 28006 Madrid, Spain. E-mail: j.g.deabajo@nanophotonics.es

The optical potential of graphene is perhaps best exemplified by the exotic properties of its collective electronic excitations, also known as plasmons. The existence of graphene plasmons has been established by the spectral signatures that they produce in the scattering of light from nanopatterned graphene ribbons (*4*). Direct spatial mapping of plasmons has also been achieved through the use of near-field optical probes (*5*, *6*). In contrast to conventional plasmon-supporting materials, such as gold, graphene must be electrically charged (that is, doped with positive or negative charge carriers) if it is to sustain low-energy plasmons. But rather than a limitation, this allows tuning of the plasmon frequencies, which increase with the density of doping charge carriers. Using microelectronic gates to supply those carriers, we can manipulate light scattered or trapped by plasmons faster than ever before (*5*, *6*) (see the figure, panel A). We thus have at our disposal a powerful tool for designing optical modulators, which currently rely on alternatives that are too slow (liquid crystals) or too bulky and expensive (nonlinear devices).

In addition to its electrical tunability, which provides a suitable tool for ultrafast electro-optics, graphene offers a remarkable versatility with respect to control of its optical properties. For example, application of strong magnetic fields leads to the formation of terahertz magnetoplasmons (*7*).



The graphene nanophotonics landscape. (**A**) Light modulation of the optical response of graphene is realized by applying voltages through electrical gates (*4*–*6*). (**B**) Molecular detection is accomplished through the modifications in the optical response associated with changes in the concentration of charge carriers (*9*). (**C**) Measurement of the electrical current modulated by photon absorption leads to efficient light detection (*11*, *12*). (**D**) Light harvesting occurs when the energy of absorbed photons is converted into charge carriers that are separated by doped gates to generate a net current.

Chemical doping is another option that has been successfully used to tune plasmons in graphene microstructures (*8*). In a related development, changes in the electronic properties of the carbon layer have been used to monitor single molecules (*9*), which could also be observed optically (see the figure, panel B).

A promising approach toward light modulation consists of decorating plasmonic nanostructures with graphene (*10*). The small changes produced in the optical response of the carbon layer upon electrical doping can be amplified by the resonant plasmonic structure. Combined with electrical detection, nanoparticle plasmons have been used to achieve nanoscale spectral photodetectors (*11*); efficient terahertz detection has also been demonstrated by combining graphene with suitably designed antennas (*12*) (see the figure, panels C and D).

The unconventional electronic band structure of graphene, characterized by a constant product of electron energy and wavelength, makes it strongly nonlinear (*13*) and provides extra functionality—the ability to control light by light. A quantum mechanical approach to this goal could exploit the extremely high optical confinement and enhancement associated with graphene plasmons (*14*), possibly leading to a robust and viable solid-state implementation of devices that can process information encoded as quantum states of light.

Advances in graphene nanophotonics are rapidly configuring a new landscape of unsuspected achievements and possibilities. The extraordinary optical nonlinearity and fast modulation of graphene are triggering novel applications in a field that is replete with challenges and opportunities. A remaining goal is to extend the spectral region of tunable graphene optical response from the infrared toward the visible and near-infrared, where it can find a larger range of applications to optical modulation, spectral light detection, and sensing. These advances will benefit from the outstanding quality of currently available graphene, although new fabrication methods need to be devised in order to retain such quality after nanostructuring and manipulation of the carbon layer. The availability of new atomically thin structures (*15*) may catalyze the extension of graphene-optics phenomena to other platforms, while opening new areas of research into the symbiotic relation between the electronic and photonic behavior in these materials.

References

1. F. Bonaccorso, Z. Sun, T. Hasan, A. C. Ferrari, *Nat. Photonics* **4**, 611 (2010).
2. A. Vakil, N. Engheta, *Science* **332**, 1291 (2011).
3. European Graphene Flagship (www.graphene-flagship.eu).
4. L. Ju *et al.*, *Nat. Nanotechnol.* **6**, 630 (2011).
5. J. Chen *et al.*, *Nature* **487**, 77 (2012).
6. Z. Fei *et al.*, *Nature* **487**, 82 (2012).
7. I. Crassee *et al.*, *Nano Lett.* **12**, 2470 (2012).
8. H. Yan *et al.*, *Nat. Nanotechnol.* **7**, 330 (2012).
9. F. Schedin *et al.*, *Nat. Mater.* **6**, 652 (2007).
10. N. K. Emani *et al.*, *Nano Lett.* **12**, 5202 (2012).
11. Z. Fang *et al.*, *Nano Lett.* **12**, 3808 (2012).
12. L. Vicarelli *et al.*, *Nat. Mater.* **11**, 865 (2012).
13. E. Hendry, P. J. Hale, J. Moger, A. K. Savchenko, S. A. Mikhailov, *Phys. Rev. Lett.* **105**, 097401 (2010).
14. F. H. L. Koppens, D. E. Chang, F. J. García de Abajo, *Nano Lett.* **11**, 3370 (2011).
15. A. N. Grigorenko, M. Polini, K. S. Novoselov, *Nat. Photonics* **6**, 749 (2012).

10.1126/science.1231119

PHYSIOLOGY

Food as a Hormone

Nutrient substrates derived from food can activate intracellular signaling cascades to regulate metabolic health.

Karen K. Ryan and Randy J. Seeley

Diet has an enormous impact on many aspects of our health, yet scientific consensus about how what we eat affects our biology remains elusive. This is especially true with respect to the ongoing debate about obesity. While many in the scientific community focus on how high-fat diets can lead to increased body weight (*1*), others assert that we should blame processed carbohydrates (*2*). Is it possible that this focus on macronutrients (i.e., fats, proteins, and sugars) is misplaced?

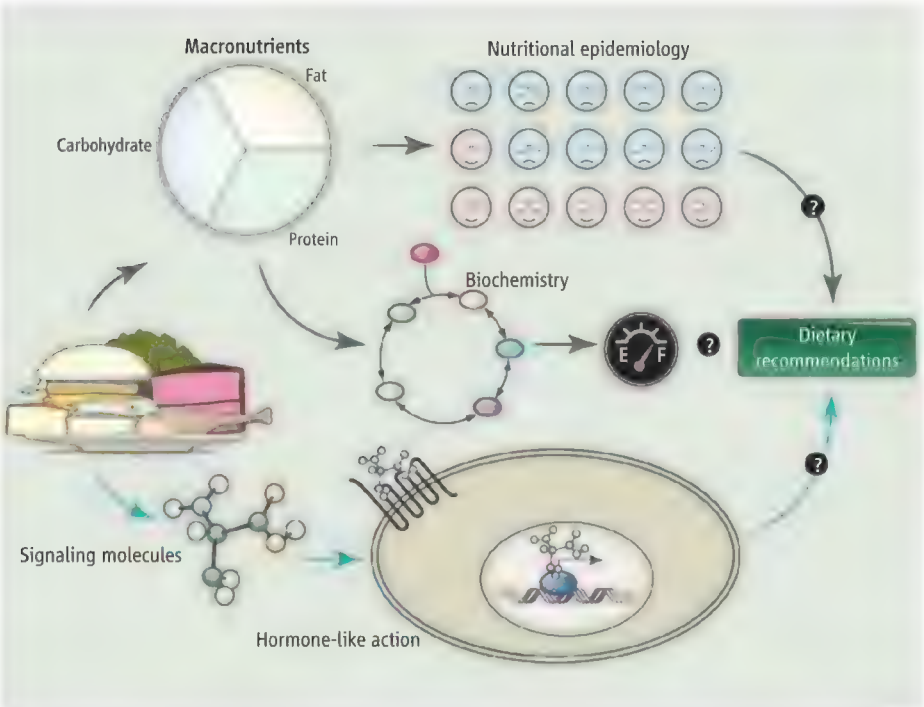
Much of the recent public discourse about the interaction between food and metabolic health relies on two basic approaches (see the figure). One is nutritional epidemiology, in which populations of people who eat different foods are compared with regard to indices of health such as body weight, with a goal of determining which diets are more or less "healthy." The other is biochemistry, in which the goal is to determine how different macronutrients are processed to yield energy. Despite valuable information provided by these two approaches, neither has resulted in a translatable scientific basis for recommending diets that improve metabolic health or reduce body weight for a large percentage of the affected population, perhaps because considering food only in terms of its macronutrient content overlooks the complexities of how food interacts with our bodies.

A growing body of evidence suggests an alternative perspective. That is, circulating substrates derived from food have specific direct and indirect actions to activate receptors and signaling pathways, in addition to providing fuel and essential micronutrients. Ultimately food can be considered as a cocktail of "hormones." A hormone is a regulatory compound produced in one organ that is transported in blood to stimulate or inhibit specific cells in another part of the body. Hormones exert their effects on target tissues by acting on cell-surface receptors to alter activity through intracellular signaling cascades or via nuclear receptors to regulate gene transcription. Although food is not produced in the body, its components travel through the blood, and nutrient substrates can act as signaling molecules by activating cell-surface or nuclear receptors.

As an example, nutritional epidemiology has touted the benefits of eating omega-3 fatty acids to protect against cardiometabolic syndrome and weight gain (*3*). Yet simple biochemistry cannot satisfactorily explain why omega-3 fatty acids should lead to benefits compared to other fatty acids. Omega-3 fatty acids bind to and activate the cell-surface receptor GPR120 (*4*), which is expressed in important metabolic tissues including adipose tissue and muscle. Reduced GPR120 signaling is associated with inflammation, weight gain, and impaired glucose control in both mice and humans (*4*, *5*). Thus, to generate the full spectrum of beneficial effects on vascular disease risk, ingested omega-3 fatty acids are not simply processed to generate energy, but additionally act via GPR120 in key tissues to improve metabolic endpoints.

Whereas activating GPR120 appears to protect against weight gain, other lipid-

Metabolic Diseases Institute, University of Cincinnati, Cincinnati, OH 45237, USA. E-mail: randy.seeley@uc.edu



Making sense of what we eat. Nutritional epidemiology and biochemical approaches, focusing primarily on the relationship between macronutrient consumption and metabolic outcomes, have not provided a translatable scientific basis to recommend diets that improve metabolic health for a broad range of people. Alternatively, understanding our diets as a collection of signaling molecules, having hormone-like actions via cell-surface and nuclear receptor signaling, may provide new insights into the relationship between what we eat and metabolic disease. Moreover, this framework may eventually allow us to make dietary recommendations from the bottom up—based on the ability of specific foods to alter relevant signaling pathways.

activated receptors exert the opposite effect. Peroxisome proliferator–activated receptor γ (PPARγ), for example, is a nuclear receptor that is activated by a variety of fatty acids and regulates transcription of genes important for lipid and glucose metabolism. Increasing PPARγ activity with pharmacological agonists enhances lipid storage in adipose tissue, and also acts in the brain to cause hyperphagia, dual actions that promote accretion of body fat (*6*–*8*). Consistent with this, reducing PPARγ activity in the brain decreases consumption of high-fat diets, thereby blunting weight gain (*6*, *8*). These studies lay the groundwork for understanding how components of high-fat diets cause overconsumption and weight gain by activating specific fatty acid receptors in the brain.

In addition to acting directly on these specialized fatty acid receptors, there is evidence that some dietary fatty acids also modify the actions of classical hormones. For example, the stomach-derived hormone ghrelin increases food intake and weight gain by binding to its receptor, growth hormone secretagogue receptor (GHSR). However, for ghrelin to signal effectively, a fatty acid must first be attached to the peptide as a side chain. Different fatty acid side chains derived from different dietary fats change the ability of ghrelin to increase food intake (*9*). These fatty acid side chains come from ingested food rather than from adipose tissue (*10*). In this way, specific dietary components can exert hormone-like metabolic effects by physical interaction with a peptide hormone.

Fatty acids are not the only direct source of "hormones" in our food; certain amino acids can also activate signaling pathways. The most-studied are the branched-chain amino acids including leucine, which activates the mammalian target of rapamycin (mTOR) pathway. mTOR is a serine-threonine kinase that regulates cell-cycle progression, growth, and insulin action (*11*). Leucine directly activates the mTOR pathway in the central nervous system to reduce food intake and body weight (*12*, *13*).

Food components also interact with gut flora to induce indirect signaling cascades within the body. For example, nondigestible complex carbohydrates, including dietary fiber, are metabolized by the gut microbiota and fermented to short-chain fatty acid (SCFA) end products—mainly acetate, propionate, and butyrate (*14*). These SCFAs bind to and activate cell-surface receptors free fatty acid receptor 2 (FFAR2) and FFAR3 to alter host metabolism. For example, FFAR2 and 3 are expressed on enteroendocrine L cells that produce the incretin hormone glucagon-like peptide–1 (GLP-1). Stimulation of L cells with SCFA induces GLP-1 secretion, but this effect is diminished in the absence of FFAR2 or, to a lesser extent, FFAR3 (*15*). Acetate and propionate also activate FFAR2 on adipocytes to increase expression of the weight-reducing hormone leptin. In this way, specific dietary carbohydrates, modified by the gut microbiota, can signal at specific receptors to alter whole-body energy and glucose metabolism.

Viewing food as a hormone could substantially influence how we make dietary recommendations to promote health or treat specific diseases. Rather than using only nutritional epidemiology to identify what healthy people consume, we may be able to design diets from the bottom up—based on their ability to alter signaling pathways in specific tissues that we know are linked to metabolic disease. In addition, this framework suggests that the argument over whether fat or sugar is to blame for the increasing incidence of obesity may be misguided. Macronutrients are classified by their energy-yielding biochemical properties, not by their ability to activate receptors in a manner similar to that of a hormone. It may be more productive to examine the signaling properties of a given diet to understand whether it will promote weight gain or weight loss. Identifying these food- and food metabolite–receptor interactions will provide new opportunities to understand the relationship between what we eat and diseases including obesity.

References and Notes

1. K. K. Ryan, S. C. Woods, R. J. Seeley, *Cell Metab.* **15**, 137 (2012).
2. R. H. Lustig, L. A. Schmidt, C. D. Brindis, *Nature* **482**, 27 (2012).
3. J. D. Buckley, P. R. C. Howe, *Nutrients* **2**, 1212 (2010).
4. D. Y. Oh *et al.*, *Cell* **142**, 687 (2010).
5. A. Ichimura *et al.*, *Nature* **483**, 350 (2012).
6. K. K. Ryan *et al.*, *Nat. Med.* **17**, 623 (2011).
7. S. Diano *et al.*, *Nat. Med.* **17**, 1121 (2011).
8. M. Lu *et al.*, *Nat. Med.* **17**, 618 (2011).
9. K. M. Heppner *et al.*, *Endocrinology* **153**, 4687 (2012).
10. H. Kirchner *et al.*, *Nat. Med.* **15**, 741 (2009).
11. S. G. Dann, G. Thomas, *FEBS Lett.* **580**, 2821 (2006).
12. D. Cota *et al.*, *Science* **312**, 927 (2006).
13. C. Blouet, H. Ono, G. J. Schwartz, *Cell Metab.* **8**, 459 (2008).
14. V. Tremaroli, F. Bäckhed, *Nature* **489**, 242 (2012).
15. G. Tolhurst *et al.*, *Diabetes* **61**, 364 (2012).

Acknowledgments: Supported by NIH (HL111319 to K.K.R.; DK093848 to R.J.S.).

10.1126/science.1234062

IBI* SERIES WINNER

Mars Student Imaging Project: Real Research by Secondary Students

The Mars Student Imaging Project, an IBI prize–winning module, provides student access to NASA scientists and spacecraft data.

Sheri Klug Boonstra† and Philip Christensen

A group of scientists gather excitedly as their image arrives from the Thermal Emission Imaging System (THEMIS) camera on NASA's Mars Odyssey orbiter. They eagerly pore over the image they targeted to answer a scientific question about the Red Planet. Working together, the team measures, analyzes, and discusses relations among surface features, as they seek to answer their research question.

Sounds like Ph.D.'s in a NASA lab, but it isn't. Instead, it's a group of secondary students in their classroom learning how science is done by conducting research using real data they acquired from Mars. Funded by NASA's Mars Public Engagement Program and the NASA Mars Exploration Program Office, the Mars Student Imaging Project (MSIP) is led by the Arizona State University (ASU) Mars Education Program under the direction of the Principal Investigator of the THEMIS camera.

MSIP began in 2002, and more than 35,000 students have participated. With a focus on Mars, which appeals to many students, MSIP is an immersive and transformational way (*1*, *2*) for students in grades 5 through early college to learn the scientific process through authentic research experiences (*3*). The project is inquiry-based and student-centered, which means that students create and investigate their own research question about the Martian surface. This project engages them in the real process of science, replacing traditional worksheets with the collection and analysis of actual data from their own, targeted image of Mars.

The MSIP curriculum evolves to stay in step with the changing needs of science classrooms (*2*). Changes are based on feedback from teachers and students, professional evaluations, emerging national education standards (specifically the Next-Generation Science Standards), and changes in technology. MSIP reflects educational best practices and integrates education research. The lessons were developed to follow the Biological Sciences Curriculum Study (BSCS) 5E inquiry instructional model and enable teachers and students without deep knowledge of planetary geology to have successful experiences (*4*). The project is flexible and connects with the traditionally taught disciplines, such as earth science, biology, and chemistry (*3*). To support educators wishing to integrate MSIP into their teaching, the ASU Mars Education Team provides online training in the lesson components tailored to meet individual needs, and team members are available to educators through virtual office hours.

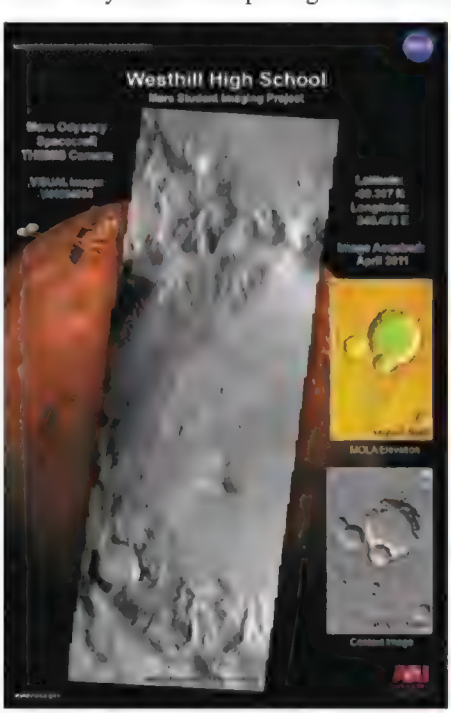
Students collaborating on Mars image analysis activity (phase 1 of the Mars Student Imaging Project).

MSIP students experience five main phases (*1*) in the project, beginning with the Mars image analysis phase (*1*, *2*, *4*). In this phase (see the photo, above), student teams work collaboratively with THEMIS images to identify common geologic features found on Mars and to develop a natural curiosity for these features and the processes that shaped them. This curiosity leads to student-generated ideas and questions shared with the entire team on topics such as wind streaks, landslides, collapsed lava tubes, dust devil tracks, lava flows, streamlined islands, chaotic terrain, channels, and dunes.

As students become engaged in the various observations, they become "student experts" in their topics of interest. They learn to make scientific observations while absorbing background information that can lead to measurable research questions (*4*). During the next phase, proposal development, each team develops a research question and articulates a testable hypothesis. This step is usually difficult because few have ever done it. The Question Mars activity was developed to aid this process. Student teams begin with a "big-picture question," such as "Why do some craters show the lobate debris aprons while others don't?"

Teachers are encouraged to help move from a general question by prompting them to develop working hypotheses. Next, the students collaborate to establish data-collection procedures, controls, and criteria. The team makes observations to determine where on Mars the features of interest occur, how measurements will be taken, what data will be recorded, and what tools will be needed.

The opportunity to struggle with these steps is crucial to inquiry learning (*1*–*3*). Students take ownership in developing their hypothesis, setting up procedures for data collection and analysis, and formulating their "why-this-is-compelling" discussion



MSIP student team's research image.

PHOTO CREDITS: (TOP) S. K. BOONSTRA; (BOTTOM) NASA/JPL-CALTECH/ASU

Arizona State University School of Earth and Space Exploration, Arizona State University, Tempe, AZ 85287–6305, USA.

*IBI, *Science* Prize for Inquiry-Based Instruction; www.sciencemag.org/site/feature/data/prizes/inquiry/.

†Author for correspondence. E-mail: sklug@asu.edu

into a formal proposal requesting THEMIS to take their image of Mars. Creating this proposal is an authentic scientific experience, as teams submit a proposal in a manner that parallels guest observers requesting time on NASA instruments.

Once the ASU Mars Education Program reviews, provides feedback, and accepts the team's proposal, the students enter the targeting phase, which consists of using the Java Mission-planning and Analysis for Remote Sensing (JMARS) tool to target (aim) the THEMIS camera at their chosen research location on Mars (see the photo, p. 920, bottom). JMARS is the same high-tech geographic information system platform used by the Mars science community to target research images and to study the martian surface. JMARS enables students to correlate their target area with other data sets in the same area taken during past or current Mars missions. Students look for surface features that connect to their research question, determine the exact location for their image, and communicate this information to the THEMIS mission planners. Student teams are given a specific number of orbits and target their image using JMARS and the actual orbit track of the Odyssey spacecraft. The student team typically assigns individual team members to collect data in a specific area pertaining to their focused research, which makes establishing criteria for data collection extremely important.

After the team receives its THEMIS image, they move on to the data analysis phase. Here, they interpret the image, collect additional data, display the data in graphical form, and draw conclusions.

As they analyze, the team is encouraged to use JMARS to obtain additional images relating to their research. The team then combines and reviews the data, identifies potential errors, and uses appropriate tools to visualize the data (e.g., graphs and tables). Teams are urged to create a wide variety of graph types to view their collected data, as this process may lead the team to discover other possible relationships in their data. Students are also asked to investigate "data outliers" that may extend the research in new directions.

Student-generated questions and research can lead to open-ended results, with many possible answers. The student team collaborates to determine the answer supported by their research and they must provide evidence from the research to support a claim. Typically, this step in the project is done in a collaborative, idea-sharing environment. The claim or evidence model is a necessary part of vetting results before the team presents the findings to Mars scientists at the symposium phase.

The project culminates when student teams interact with ASU Mars scientists during the final report and symposium phase. Because MSIP is offered in two formats, on-site and distance learning, students across the United States are able to participate in this phase. The on-site option is a 2-day program in which student teams complete their proposal and targeting at their own school. They then travel to ASU to work with Mars Education staff and scientists to finish the research and present their final project in a symposium to several Mars scientists face-to-face. The distance learning option involves a minimum of two live sessions: an online meeting to present the proposal and a final virtual meeting for the symposium with Mars scientists. Scientists involved in MSIP help the teams by vetting the students' approach and data. They probe the students' reasoning and discuss their findings. Some MSIP student projects have resulted in scientific discoveries (*5*), technical publications (*6*), and poster sessions at technical conferences (*7*).

MSIP provides an authentic research experience where students have access to real data sets from which many different questions can be asked. Working from a question they generate themselves, MSIP participants learn that, in science, it is the evidence and defensible interpretation through discourse that will determine the answer. Although the participating students are not Ph.D.'s in a NASA lab, they become scientists, nevertheless.

PHOTO CREDIT: (LEFT) J. ANTHONY ZIPPAY/ASU (RIGHT) TOM STORY/ASU

About the authors

Sheri Klug Boonstra is the director of the ASU Mars Education Program and the formal education lead for NASA's Mars Public Engagement Program. She is the precollege representative for NASA's Science Mission Directorate Planetary Science Education and Public Outreach Forum and served as the Education and Public Outreach representative on the Solar System Exploration Subcommittee for NASA Headquarters. She was the Universities Space Research Association principal investigator for NASA's Undergraduate Student Research Program and has experience as a classroom science teacher. Ms. Klug Boonstra has received several NASA group achievement awards for her participation in NASA missions and programs. **Philip Christensen** is a Regents' Professor of Geological Sciences in the School of Earth and Space Exploration at ASU. He is the principal investigator for THEMIS on NASA's Mars Odyssey orbiter. He is developing the OSIRIS-REx Thermal Emission Spectrometer for NASA's OSIRIS-REx asteroid sample-return mission, to launch 2016. He received NASA's Exceptional Scientific Achievement Medal in 2003 and its Public Service Medal in 2005. He was elected Fellow of the American Geophysical Union in 2004 and received the Geological Society of America's G. K. Gilbert Award in 2008.

References and Notes

1. S. R. Singer, M. L. Hilton, H. A. Schweingruber, Eds., *America's Lab Report: Investigations in High School Science* (National Research Council, National Academies Press, Washington, DC, 2006).
2. R. W. Bybee *et al.*, *The BSCS 5e Instructional Model: Origins and Effectiveness* (Biological Sciences Curriculum Study, Colorado Springs, CO, 2006).
3. R. W. Bybee *et al.*, The BSCS 5e instructional model and 21st century skills (National Academies of Science, Washington, DC, 2009); www7.nationalacademies.org/bose/1Bybee_21st%20Century_Paper.pdf.
4. J. D. Bransford, A. L. Brown, R. R. Cocking, Eds., *How People Learn: Brain, Mind, Experience, and School* (National Academies Press, Washington, DC, 2000).
5. R. Burnham, *ASU News Now*, 17 June 2010; https://asunews.asu.edu/20100617_skylight.
6. R. Chan *et al.*, presentation, 12th Australian Space Science Conference, Melbourne, Australia, 24 to 26 September 2012.
7. D. Nguyen *et al.*, 36th Lunar and Planetary Science Conference, League City, Texas, 14 to 18 March 2005, Abstr. no. 2036.

Acknowledgments: This work was made possible with the creative collaboration and funding support of M. Viotti of the Mars Public Engagement Program, NASA's Mars Odyssey Project, and NASA's Mars Program Office. C. Bowman provided valuable evaluation data and insights. We thank G. Swenson and our partner, the Peoria Unified School District, for ongoing project support and willingness to test new iterations of MSIP. The ASU MSIP Team is recognized for all their design and implementation efforts. R. Burnham, D. Boonstra, J. Swann, J. A. Zippay, and L. Manfredi provided research and editorial support. Finally, a special thanks to all the Mars researchers, MSIP teachers, students, and staff who have contributed to the growth and impact of the Mars Student Imaging Project during the past decade.

Supplementary Materials
www.sciencemag.org/cgi/content/full/339/6122/920/DC1

10.1126/science.1229849

ASSOCIATION AFFAIRS

Phillip A. Sharp: Supporting Science and Engineering as Innovative Forces

Phillip Sharp takes his responsibilities as a role model seriously, from speaking in schools about his childhood interest in science on a Kentucky farm, to posing with rock stars in the pages of *GQ* to raise interest in cancer research.

The magazine spread has brought him more attention than his 1993 Nobel Prize, "but I was pleased to do it," he said during a recent interview. "I think any time scientists are presented as being engaged in society, understanding what's going on in society, and in a positive mode, it's a big plus for science."

Sharp's career has followed a trajectory that might itself serve as a model for the future of science, with an emphasis on interdisciplinary research, innovation, and the translation of scientific discovery. As the MIT molecular biologist and biochemist assumes the AAAS presidency, he sees a prominent leadership role for the organization in these areas.

In particular, Sharp said, "AAAS will need to be at the center of the debate on R&D support in the next few years, and I am ready to help play a role as a science and engineering advocate." He praised AAAS for the work it has already done to highlight the connections between science funding and global economic growth. "We cannot let that slip, because if we do, we're going to pay the price for the next 30 years," he said. "And every person in this country is going to pay the price, because the jobs aren't going to be there."

Science funding will be a challenge "in almost every developed and developing country" as well, Sharp said. AAAS could play a vital role in addressing some of these challenges, he noted, drawing on the strong collaborations it has established with international scientists and scientific societies. "In a world that is increasingly 'flat,' where science, technology, and economies in each community are interdependent," he suggested, "the importance of AAAS is greater today than ever before."

Support for interdisciplinary science will also be key, he said, citing its potential to propel innovation and open up entirely new fields of research. Sharp has written extensively about convergence—where the life sciences meet physical and computational sciences and engineering—as co-chair of the U.S. National Academy of Sciences committee that produced the report, *A New Biology for the 21st Century,* and in other white papers.

Phillip A. Sharp

Sharp said that universities, industries, and organizations like AAAS "have moved a long way" toward creating environments where convergence thrives. "We're making changes as a community that make it easier for an individual to be engaged in interdisciplinary science, and we're encouraging exciting new areas of science to develop as well."

After winning the Nobel Prize for his co-discovery of "split genes," demonstrating that genes can be discontinuous on a strand of DNA, Sharp has seen the power of convergence firsthand in his research on cancer cell biology. He has moved deeper into explorations of RNA interference, or RNAi, in which RNA molecules can be used as "on/off switches" for gene expression. This powerful tool, along with advances in computational power and whole genome sequencing, "have made this a wonderful moment in science."

Sharp was a convergence pioneer of another sort in the 1970s when he co-founded Biogen (now Biogen Idec), helping to shape the future of a nascent biotechnology industry. "This was a time when technology and genetic engineering and recombinant DNA were newly emerging in the lab, and it was widely recognized and clear to me that there would be great benefit to society in having this knowledge translated." In 2002 he continued those steps toward translation as a co-founder of Alnylam Pharmaceuticals, an early-stage RNAi therapeutics company. Working with both companies, Sharp said, "was a way of contributing to society and learning at the same time that I found very enriching."

In a way that is perhaps fitting for the son of farmers and ironworkers, Sharp emphasizes the importance of engineers in sharing the benefits of science. The United States produces too few engineers, he said, and "they have a skill set that is essential for our media, for life science, for medicine, and for manufacturing. We need to expand this if we're going to have a knowledge and work base in this country that can actually deal with the innovation we need to keep our economy growing."

Sharp is an Institute Professor at the Koch Institute for Integrative Cancer Research at MIT, where he served as director from 1985 to 1991. He also served as the founding director of MIT's McGovern Institute for Brain Research from 2000 to 2004. Sharp succeeded William H. Press, a professor of computer science and integrative biology at the University of Texas at Austin, at the close of the AAAS Annual Meeting on 18 February. Press is now serving a 1-year term as chairperson of the AAAS Board.

Along the Licking River in Falmouth, Kentucky, Sharp raised cattle and grew tobacco to pay for his college education. He encourages more engineers and scientists to talk with rural and poor students, especially, as a way of inspiring the next generation of researchers. In some of these communities, Sharp said, students "don't see people who have made contributions to science and engineering. Someone standing in front of them and telling them how exciting it is can have a profound impact on a young mind."

"Every student who wants to contribute and has the opportunity should think about" a career in the sciences, he said. "The growing importance of science and engineering in society, in terms of helping people, is a way of contributing to everyone across the face of the earth."

–Becky Ham

CREDIT: PHOTO COURTESY PHILLIP A. SHARP

How Cells Know Where They Are

Arthur D. Lander

Development, regeneration, and even day-to-day physiology require plant and animal cells to make decisions based on their locations. The principles by which cells may do this are deceptively straightforward. But when reliability needs to be high—as often occurs during development—successful strategies tend to be anything but simple. Increasingly, the challenge facing biologists is to relate the diverse diffusible molecules, control circuits, and gene regulatory networks that help cells know where they are to the varied, sometimes stringent, constraints imposed by the need for real-world precision and accuracy.

To measure the distance from one object to another, something must traverse the intervening space. For distances of centimeters or meters, we lay down a ruler or tape measure; for objects much farther away, we might bounce off sound or radio waves and measure the return time. In the microscopic world of cells, where measurements of position typically are made over spans of a few hundred micrometers or less (Fig. 1), the things that most readily cross such distances are molecules, and the simplest way they do so is by aqueous diffusion. Unlike the propagation of light or sound waves, diffusion is not a constant-rate phenomenon—a diffusing front gets slower as it spreads—making computation of distance from arrival time tricky (although not impossible). But, given constant production at a source, diffusion can create steady-state gradients within which concentration is a proxy for distance. From this insight, it was proposed, and later demonstrated, that cells in developing animal embryos receive positional cues from diffusible molecules that indeed form stable gradients across tissues (*1*). Such molecules, dubbed morphogens, play central roles in orchestrating developmental pattern formation.

Recently, there has been debate about whether cells really receive positional information by measuring concentrations in steady-state diffusion gradients (*2*, *3*). The most serious objections have to do with reliability: In the world of cellular biochemistry, variability in synthesis and secretion, in the binding of molecules to receptors, in the activation of signaling pathways, and in gene regulation can all be quite high. Among other things, such variability can stem from the environment (e.g., unpredictable temperature or nutrition), genetics, or stochastic fluctuations in biochemical processes. Yet the positional information that cells ultimately obtain is often exceedingly reliable, particularly during development (as evidenced by the remarkably accurate symmetries and family resemblances we encounter in our own bodies). Can steady-state diffusion gradients provide that kind of reliability?

Department of Developmental and Cell Biology, Department of Biomedical Engineering, and Center for Complex Biological Systems; University of California Irvine, Irvine, CA 92697, USA. E-mail: adlander@uci.edu

In short, it depends. It depends on the amount and kind of variability cells face, the mechanisms by which gradients form, and how much reliability is required. Consider, for example, a sheet of cells (an epithelium) in which a diffusible morphogen is secreted at a constant rate by cells lying in a stripe (Fig. 2A). The morphogen is destroyed everywhere, through receptor-mediated uptake, at a constant proportion per time (this situation approximates what is thought to be the case in a variety of developing tissues). Eventually, a stable gradient forms in which morphogen concentration falls exponentially away from the source, the precise shape determined by the morphogen's rate of production, diffusivity, and rate of uptake and destruction. A cell's reading of morphogen concentration will then depend on its number of receptors and how much intracellular signaling occurs per occupied receptor.

Not surprisingly, if cellular location is measured from the morphogen concentration sensed by each cell, unreliability in any of these processes—morphogen production, transport, uptake, receptor synthesis, and signaling—will produce measurement errors. The type and magnitude of the error will depend on what is varying and where the cell is located. Variability that enters upstream of individual cells (e.g., in morphogen production or transport) or affects all cells equally (e.g., animal-to-animal differences or temperature change) will produce inaccuracy, that is, a shift in the locations of positional values. In contrast, cell-to-cell variability produces imprecision, that is, scatter in the

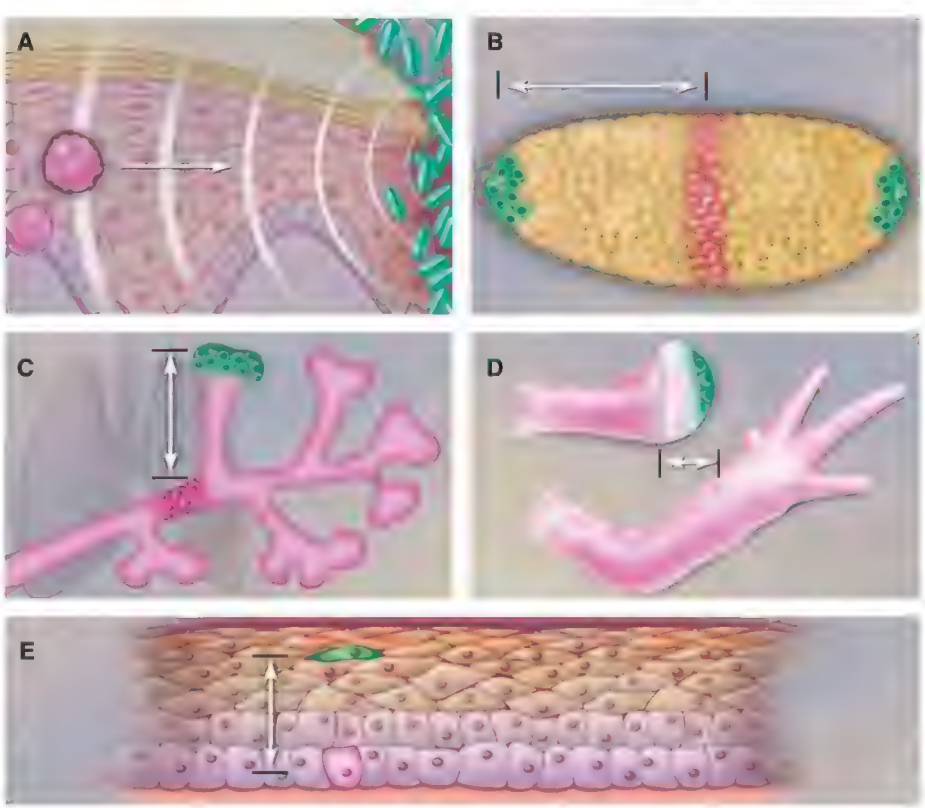


Fig. 1. A gallery of positional tasks. A leukocyte (**A**) may need to know in which direction to head to find the site of an infection but not the absolute distance to it. A cell in an early embryo (**B**) may need to know absolute location with respect to one or the other end of the embryo so that it differentiates into a spatially appropriate cell type, whereas a cell in a tissue undergoing branching morphogenesis (**C**) may need to know only the rough location with respect to the nearest branch point or vessel. In regenerating tissues (**D**), cells need to know their position with respect to a site of injury or amputation, whereas in tissues or organs with laminar structures (**E**) cells may need only know whether they are in the appropriate layer.

positional measurements that cells at equivalent locations make. Inaccuracy is often best quantified by a sensitivity coefficient, a unitless number that captures the fold change in the location of a positional value per any given fold change in some upstream process. Imprecision is most readily quantified as a transition width, the distance away from any position x that one must move before no more than some maximum proportion of cells (e.g., 15%) conclude that they are on the wrong side of x from where they actually are (*4*–*6*).

Sensitivity coefficients for perturbations of morphogen and receptor production and estimated transition widths associated with stochastic fluctuations in receptor occupancy are plotted in Fig. 2B for morphogen gradients such as the one in Fig. 2A by using assumptions about receptor abundance estimated from the *Drosophila* wing disc (*6*). Sensitivity to morphogen production declines with distance from the morphogen source, whereas sensitivity to receptor production falls to zero and then rises to an asymptotic value. The transition width caused by fluctuation in receptor occupancy ("binding noise") rises with distance from the morphogen source, because lower receptor occupancy produces larger stochastic fluctuations (*6*). A similar rise occurs for fluctuations driven by other sources of cell-to-cell variability (*5*).

In essence, there is some effect of every form of variability at virtually every location, but which errors matter in practice depend on where positional measurements need to be made, how much error can be tolerated, and how much different inputs actually vary. Only in a few cases are hard data on in vivo variability available [e.g., (*5*, *7*, *8*)], but the fact that mutants that are heterozygous null for genes encoding morphogens, receptors, or components of signaling pathways (i.e., putative 50% reductions in input) usually display developmental patterns that are shifted only slightly (e.g., <20%) suggests that sensitivity coefficients for developmental morphogen gradients are often <0.26 (ln1.2/ln2) (*6*). Similarly, from looking at the phenotypic effects of mutations that lead to broadening of the widths of gene expression stripes established by morphogen gradients [e.g., (*9*)], we can estimate bounds on tolerable transitions widths (Fig. 2B). Such limits reveal a surprising fact: In simple, steady-state morphogen diffusion gradients, there is no location at which cells simultaneously achieve substantial robustness to variation in morphogen production, variation in receptor production, and cell-to-cell variability in receptor occupancy.

What Price Reliability?

Might cells improve robustness by altering the mechanisms by which gradients form or are sensed? Evidence suggests that they do. For example, some cells increase morphogen degradation in response to morphogen signaling (so-called self-enhanced degradation), which decreases sensitivity to variation in morphogen production (*10*). But this benefit comes at the expense of making gradients shallower, the result of which is greater imprecision for any given amount of cell-to-cell variability (*6*). Of course, precision can be improved by operating in regimes of higher total morphogen (increased receptor occupancy translates into reduced stochastic fluctuations in binding), but this quickly leads to receptor saturation near the morphogen source, which greatly increases sensitivity to variation in morphogen production rate (*6*). Sensitivity to morphogen production rate can be lowered by measuring morphogen concentrations before the gradient reaches a steady state (*2*), but this carries the price of making cells highly dependent on the precise time at which they make their measurement (i.e., they gain a more accurate ruler at the expense of needing an accurate clock).

These scenarios illustrate a basic engineering principle: Strategies that improve performance in one arena typically degrade it in another. Often such trade-offs arise out of a linkage between the number of degrees of freedom available for combating unreliability and the number of sources of unreliability. For example, assume that we could insert into the morphogen gradient in Fig. 2 an additional component, a "nonreceptor" molecule that binds, internalizes, and degrades the morphogen without participating in signaling [such a role has been suggested for cell-surface proteoglycans (*11*, *12*)]. Now the curve of sensitivity of positional measurement to receptor production can easily be shifted so that it overlaps with the curve of sensitivity to morphogen production (i.e., the system can now be simultaneously robust, at many locations, to variation in both morphogen and receptor production). Unfortunately, addition of a new component adds a new potential source of unreliability—variability in the production of the nonreceptor. Not surprisingly, just those conditions that shift the receptor-sensitivity curve to a more desirable location correspond to conditions that make nonreceptor sensitivity high everywhere. In other words, whenever we add something new, we risk trading sensitivity to what was there for sensitivity to what we added.

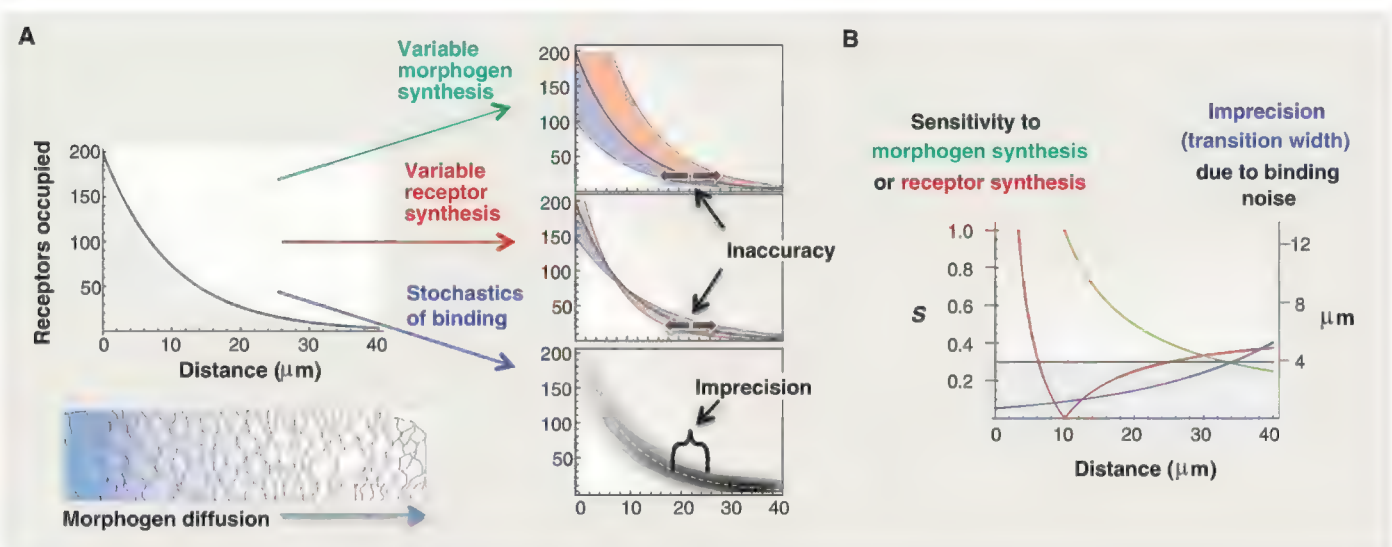


Fig. 2. Effect of input variability on the reliability of diffusion gradients. **(A)** Diffusion of molecules through intracellular spaces, when coupled to receptor-mediated uptake, produces steady-state gradients from which cells can ascertain their positions. But variability in processes that contribute to gradient formation or interpretation will necessarily lead cells to make mistakes. Their errors may be classified as either inaccuracy, whereby the average cell at a given location obtains an incorrect positional value, or imprecision, whereby there is cell-to-cell variability in the positional information obtained by cells at equivalent positions (the latter effectively converts the gradient of positional information into a probability cloud, rather than a sharp curve). **(B)** The most important potential sources of unreliability are different at different locations along a gradient [values shown are based on the gradient in (A)].

The same problem surfaces when we examine the suggestion that morphogen gradients improve accuracy by abandoning simple morphogen diffusion in favor of active transport (*13*). Sensitivity to variation in morphogen production can certainly be reduced but only at the expense of creating new sensitivity to variation in the level of whatever carrier is responsible for mediating the active transport (e.g., receptors). In fact, observations suggesting that active transport plays a role in morphogen-gradient formation in animal tissues (*14*–*16*) have increasingly been challenged (*17*–*20*). The situation is clearly different in plants, which lack the contiguous intercellular spaces of animal tissues, so that long-range transport of water-soluble molecules typically necessitates passage through cells, enabled by plasma membrane carriers and active transporters (*21*, *22*).

Of course, relying on diffusion for transport does make morphogen gradients highly sensitive to variations in diffusivity, but the beauty of diffusion is its reliability. Diffusion coefficients are only weakly sensitive to temperature and are independent of the metabolic state of the cell. Effective diffusivity is surprisingly insensitive to randomly placed obstacles, even when densely packed (*23*), and transient binding or trapping by "sticky" molecules in the environment only affects the rate at which morphogen gradients form, not their steady-state shape (*2*, *18*, *19*).

This discussion highlights an important point: The goal of good design is not to make sensitivity go away—because of performance trade-offs every system will always be sensitive to something—but rather to shift it onto the things that are reliable and away from those that are not. Of course what is reliable may be idiosyncratic. For example, tissues with small cells that have few receptors will experience larger stochastic fluctuations in receptor occupancy than tissues with larger cells and many receptors, placing a greater premium on insensitivity to receptor density in one case versus the other. Likewise, embryos that develop in eggs laid on land should place a greater premium on resistance to temperature fluctuation than mammalian embryos or embryos that develop in marine environments.

An interesting example is provided by the retinoic acid (RA) gradient that patterns the vertebrate hindbrain. Because RA is a small molecule derived in a few enzymatic steps from dietary vitamin A, we expect its abundance to vary greatly from embryo to embryo, much more than is the case for morphogens encoded by the genome. It should thus come as no surprise that numerous control mechanisms that reduce or counteract such variability—including feedback regulation of RA biosynthesis, degradation, and delivery to intracellular receptors—are observed in the hindbrain (*24*). Consistent with the presence of these (and other) control mechanisms, large experimentally induced fluctuations (>10-fold) in total RA levels have surprisingly little effect on the positional information that hindbrain cells obtain from RA (*25*). From the standpoint of the organism, a reduced sensitivity to variation in abundance of RA undoubtedly increases sensitivity to amounts of other components [e.g., cellular RA-binding proteins and cytochrome P450 and RA receptors (*26*, *27*)]. This is an acceptable trade, because amounts of those other components are likely to be more reliable.

Quality Through Quantity

Shifting sensitivities is not the only way to improve positional reliability; by making combinations of measurements, cells may exploit strategies such as pooled sampling and disturbance compensation. The former refers to the way scientists typically tame noisy data: We average (pool) repeated measurements. For this to be effective, successive measurements must be independent; that is, disturbances that corrupt one measurement need to be uncorrelated with those that corrupt the next.

Cells naturally pool measurements by letting intracellular signals accumulate over time (temporal integration). For example, morphogen-receptor complexes may be internalized and continue to signal within endosomes. Or the half-lives of protein products produced in response to morphogen signaling may exceed the duration of signaling complexes. In such cases, steady-state morphogen responses will reflect the average morphogen concentration over a time on the order of the half-life of signaling endosomes or downstream protein products, whichever is greater.

In this way, many short-lived disturbances are easily averaged out (for example, fluctuations in the extracellular levels of freely diffusing molecules tend to relax on a time scale of seconds). But some disturbances may have long time scales. For example, stochastic fluctuation in receptor occupancy resulting from the probabilistic nature of binding will reflect the time scales of receptor dynamics, which generally need to be long (e.g., hours) so as not to interfere with the formation of long-range diffusion gradients (*6*, *28*). Whether cells have sufficient time to average away such noise may thus depend on how quickly they need to obtain their positional information. In other words, cells face a speed-accuracy trade-off (Fig. 3).

A drawback of simple temporal integration is that a cell must commit to a particular time scale of integration (e.g., a particular half-life of the accumulating molecule). Even if disturbances happen on faster time scales, cells must wait a fixed time before determining their locations [lest their measurements be inaccurate (Fig. 3C)]. One way to get around this is to exploit noise-induced switching (*29*), a phenomenon that occurs when noisy inputs act on hysteretic switches—devices that switch from off to on at different thresholds from which they switch from on to off. Between the two thresholds, any transient fluctuation that flips the switch from off to on will tend to leave it in the on state until a large enough fluctuation in the other direction occurs to flip it back. In effect, the on state is "remembered" for a time related to the rate at which fluctuations occur. Because of this memory, hysteretic switches achieve temporal averaging on a time scale set by the noise itself (Fig. 3D).

There are other reasons why switches are a useful thing to include in the machinery with which cells read position: Diffusible carriers of positional information (morphogens) are usually smoothly graded in space, but cells often need to make binary decisions (e.g., to differentiate or not). Response circuits that are switchlike, that is, ultrasensitive, are thus essential. What is interesting is how many of the switches that operate downstream of positional cues turn out to be of the hysteretic type. Examples include gene regulatory switches driven by Bicoid signaling in the *Drosophila* embryo; sonic hedgehog, fibroblast growth factors (FGFs), and RA in the vertebrate spinal cord; and RA in the vertebrate hindbrain (*30*–*34*). In several of these cases, modeling strongly suggests that noise-induced switching plays a key role in producing sharp, positionally accurate gene expression borders (*30*, *34*).

Pooled sampling does not necessarily mean repeating the same measurement; it can also mean pooling different kinds of data. For example, some cells can get positional information from more than one morphogen at the same time. In the vertebrate hindbrain, the posterior-to-anterior RA gradient is supplemented by FGF and Wnt gradients in the same orientation (*35*); in the early *Drosophila* embryo, the Bicoid gradient works together with independent gradients of Caudal and maternal Hunchback (*7*); in the bone morphogenetic protein gradients that provide dorsoventral positional information to early invertebrate and vertebrate embryos, there are always multiple ligands (*36*, *37*).

Measurements can also be pooled over space instead of time; that is, neighboring cells can share information. In syncytia, such as the early *Drosophila* embryo, some spatial integration happens simply by virtue of diffusion of downstream effectors of morphogen signaling from one nucleus to another (*38*). In cellularized systems, mechanisms of spatial pooling are less well understood. One fascinating example occurs in the *Drosophila* wing imaginal disc, where the spatial extent of expression of the transcription factor *vestigial* (*vg*, which directs cells to adopt a wing fate) is determined by the Wingless (Wg) morphogen gradient. Because this gradient is quite steep, Wg concentrations are thought to be very low at the edges of the *vg* domain, exposing receptor occupancy to large, slow, stochastic fluctuations. Ordinarily, this should make the formation of a sharp, reliable gene expression border very difficult, but cells communicate with their neighbors (via the Fat signaling pathway) so that Wg is only permitted to turn on *vg* in a cell adjacent to one already expressing *vg* (an example of spatial pooling). Moreover, once they turn it on it remains on independent of Wg (i.e., the switch is hysteretic, so they accomplish temporal integration as well) (*39*). The combined effect is to

produce a wave of *vg* expression that spreads from the Wg source outward, coming to a near halt at a reliable position (Fig. 3D, curve g).

Fixing Errors Before They Happen

Making multiple measurements that are independent with respect to disturbances is not always feasible. In such cases, cells may exploit a different strategy, disturbance compensation. This becomes available whenever multiple measurements are affected by a common disturbance and there is a priori knowledge of the nature of their mutual dependence. A particularly simple example is ratiometric measurement: For instance, if a morphogen's rate of production rises in a certain way with temperature, then measuring its abundance relative to that of some other molecule that rises in the same way with temperature will produce a temperature-corrected reading.

Ratiometric measurement can be implemented in a surprising number of ways. For example, cell-to-cell variation in receptor occupancy resulting from noisy receptor expression can be nullified by measuring the ratio of occupied to unoccupied receptors rather than receptor occupancy per se. Indeed, amounts of the morphogen Hedgehog appear to be measured in just this way (*40*). Similarly, when morphogens with opposing actions are produced at opposite ends of a field of cells, the net signal that cells receive can cancel out perturbations that affect both morphogens equally (*41*).

Any instance in which perturbations to a system have more than one predictable effect creates an opportunity for disturbance compensation. For example, in a morphogen gradient set up by

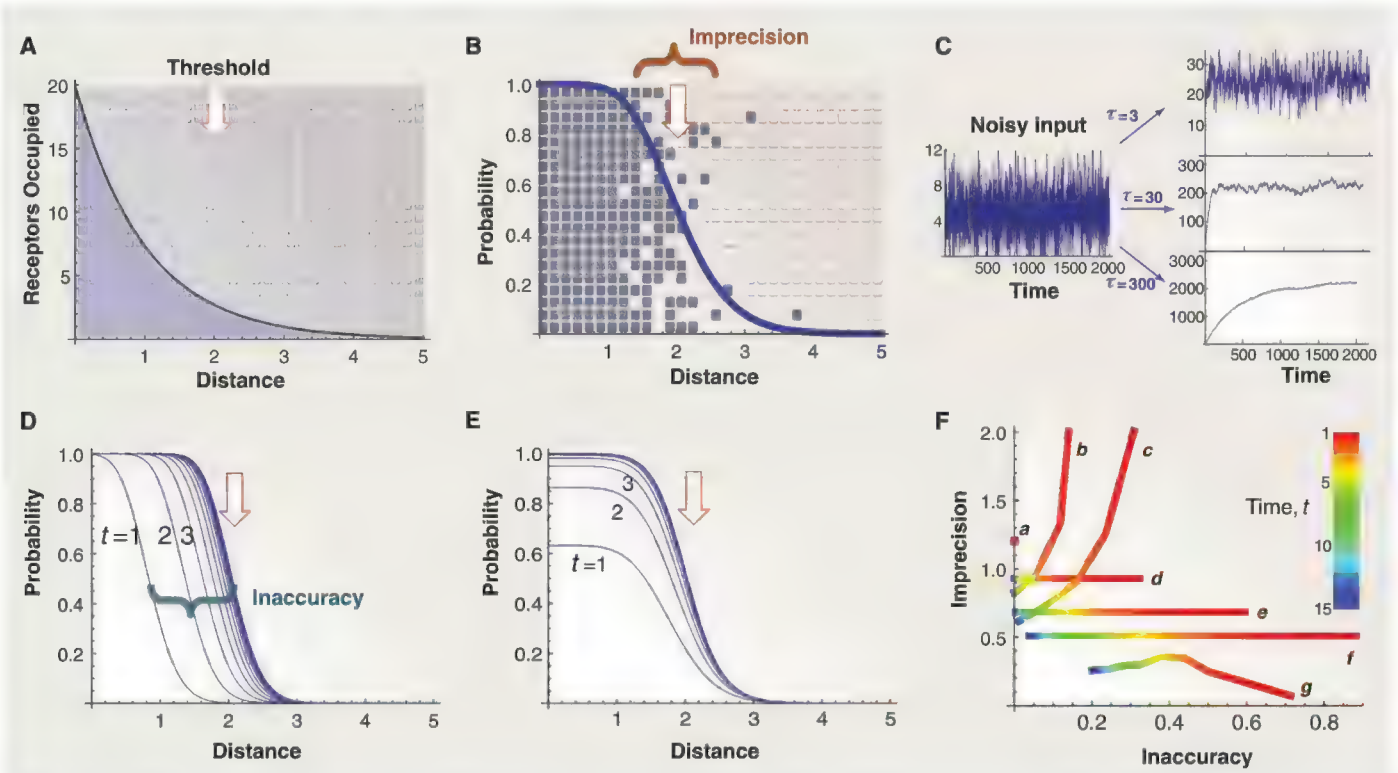


Fig. 3. The costs of pooling information. Even when average receptor occupancy accurately reflects position (**A**), any single measurement made by a cell will be corrupted by detection noise, including (but not limited to) stochastic variation in receptor occupancy. (**B**) Dark squares correspond to cells with instantaneous receptor occupancy above an arbitrary threshold (here equal to 2), and the curve depicts, given just the sampling noise in the receptor occupancy levels in (A), the probability of individual cells exceeding this threshold as a function of distance. Imprecision is the zone over which a cell's measurement has a substantial probability of being on the "wrong" side of the threshold value, given its location. (**C**) A common form of temporal integration occurs when noisy signals drive the accumulation of a downstream element (e.g., a second messenger or a transcription factor) that turns over more slowly than the noise fluctuations. The simulation illustrates what happens when a signal with random (Gaussian) noise drives accumulation of a factor with a half-life of 3, 30, or 300 times the characteristic noise time. The longer the integration, the lower the noise, but the longer it takes for the signal to approach the correct steady-state value. (**D**) Trade-offs between noise reduction and speed necessarily also apply to cells that read morphogen concentrations. In this simulation, noisy signaling like that in (B) drives production of a molecule with a half-life twice the characteristic time t of fluctuations in receptor occupancy [axes as in (B)]. Imprecision is reduced by ~45%, but only after sufficient integration time ($t > 4$) is the threshold value of morphogen signaling crossed near its steady-state location. (**E**) Here, pooling was achieved by noise-induced switching, in which hysteretic cellular decisions drive cells back and forth across threshold values in a manner driven by the noise itself. (**F**) The trade-offs between precision, accuracy, and timing are different for different strategies. Several simulations like those in (D) and (E) are summarized by using curves (a to g) to plot the time evolution of imprecision and inaccuracy, with color used to depict time (that is, simulations start at red and end at blue; time is expressed in units of the characteristic noise time, i.e., the time it takes noise to decorrelate). Imprecision is in units of transition width (normalized to morphogen gradient length scale), and accuracy as the fractional approach of the location of median probability to that at $t = \infty$ (b to f) or $t = 100$ (g). Without any pooling (a), imprecision is high, but reads are accurate from the outset. With noise-induced switching (b and c), precision is initially very poor but eventually exceeds that in a; the detailed trajectories depend on whether the hysteretic window is large (0 to 4 in b); or more modest (1 to 3 in c). With temporal integration (d to f), precision improves steadily with longer integration times (d, e, and f depict $\tau = 1$, 2, and 4, respectively), but for dramatic improvement it can take very long to achieve accuracy. Curve g simulates the spatiotemporal strategy used by the fruit fly wing disc to measure the Wg gradient (see text); here very high precision can be achieved, with inaccuracy that declines at a rate determined by how fast information about *vg* expression is relayed from cell to cell (in the simulation used to generate the figure, a rate of spread of six cells per time unit was used).

diffusion and receptor-mediated uptake, changes in receptor production do not just alter the amount of morphogen detected by each cell (as depicted by the red curve in Fig. 2B), they also alter the slope of the gradient. If cells can directly measure that slope, its change can be used to correct measurements of position obtained from morphogen levels. In fact, cells in the *Drosophila* wing disc do measure the slope of the Decapentaplegic (Dpp) morphogen gradient, albeit indirectly, because the Fat signaling pathway responds to cell-to-cell differences in perceived morphogen (*42*). Whether this information is actually exploited to correct for disc-to-disc variations in receptor levels remains to be seen.

Even the noise in a signal can drive disturbance compensation, provided there is a known correlation between noise strength and signal strength (generally true for sampling noise, as in the biochemical fluctuations in ligand-receptor interactions). That is because hysteretic switches not only time-integrate in response to noise, they tend to shift their input-output relationships as a function of the noise strength (*29*). Indeed, some of the improvement in precision provided by noise-induced switching (Fig. 3D, curves b and c) reflects this effect.

Measurement Through Self-Organization

In the earlier discussion of how *vg* expression is controlled in the *Drosophila* wing disc, the morphogen Wg was presented as the source of positional information, with cell-cell interactions serving to help cells read Wg more reliably. We could, however, flip this depiction on its head and argue that the cell-cell interactions themselves carry the positional information. That is because the arrival of a spreading wave of cell-to-cell signaling [a "juxtacrine relay" (*43*)] can itself be used to measure distances in much the same way that the time of arrival of thunder allows us to judge distance from a lightening strike. From this viewpoint, we might see the Wg gradient as merely biasing the rate at which the wave travels so as to improve the accuracy with which position is determined from it.

There are, in fact, many ways in which short-range cell-cell interactions can make long-range things happen reliably. Systems that do this are called spatially self-organizing, because spatial order emerges directly from collective, or pooled, interactions. Both the Fat and the Notch pathways have the potential to drive self-organization based on interactions between cells and their immediate neighbors (*22*, *42*, *44*). Self-organization over longer ranges can be achieved through the local production of secondary morphogens and inhibitors that may interfere to create spontaneous patterns (e.g., Turing patterns) with characteristic length scales (*45*). When allowances are made for cell rearrangement, even differences in cell-cell adhesion will drive spontaneous, large-scale spatial organization (*46*).

Of course self-organization cannot tell cells where they are relative to a fixed reference point unless that location is somehow linked to the self-organization process. Such coupling can come from boundary conditions [as occurs when bacterial cells use self-organization to locate their own midpoints (*47*)], but the control of *vg* expression in the wing disc tells us that in tissues it can also come directly from a long-range morphogen gradient. These observations suggest that we risk being narrow-minded when we think of cells as first measuring their positions and then acting in a position-specific way (e.g., organizing into patterns). We may find that, more often than not, the spatial control of morphogenesis is inextricable from the process of morphogenesis itself.

On the Horizon

In biology, simple questions rarely have simple answers, and "how do cells know where they are?" is no exception. I have focused here on the problem that, despite the existence of straightforward ways for cells to measure position, making measurements sufficiently accurate and precise is inherently challenging. Although recent years have seen considerable progress in identifying mechanisms for encoding and detecting positional information, as well as for achieving reliability, there is much we still do not know. For example:

1) How many of the "rulers" that cells use have we found? Are they mainly diffusing molecules, or are migrating cells, mechanical signals, or even electrical fields just as important?

2) How collaborative are most cases of positional sensing? As we have seen, cells gain reliability by pooling information with neighbors, or even participating in large-scale, spatial self-organization. How often do they do so, and in what contexts?

3) What are the primary sources of unreliability that constrain how cells in different contexts measure position? To answer this question requires that we not only quantify natural variability but also learn about the environments in which organisms evolved.

4) To what extent can known sources of uncertainty and variability provide a satisfactory (constraining) explanation for position-sensing mechanisms we observe? In particular, how many of the multiple morphogens, feedback loops, and complex gene regulatory circuits that we find in real systems can we explain by their positive influence on reliability?

Tackling these questions may force us to explore new research directions, but if we wish to truly understand how cells read position, there is really no better position to be in.

References and Notes

1. A. D. Lander, *Cell* **128**, 245 (2007).
2. S. Bergmann *et al.*, *PLoS Biol.* **5**, e46 (2007).
3. M. Nahmad, A. D. Lander, *Curr. Opin. Genet. Dev.* **21**, 726 (2011).
4. F. Tostevin, P. R. ten Wolde, M. Howard, *PLOS Comput. Biol.* **3**, e78 (2007).
5. T. Bollenbach *et al.*, *Development* **135**, 1137 (2008).
6. A. D. Lander, W. C. Lo, Q. Nie, F. Y. Wan, *Cold Spring Harb. Perspect. Biol.* **1**, a002022 (2009).
7. V. V. Gursky *et al.*, *BMC Syst. Biol.* **5**, 118 (2011).
8. D. M. Holloway *et al.*, *PLOS Comput. Biol.* **7**, e1001069 (2011).
9. B. Biehs, M. A. Sturtevant, E. Bier, *Development* **125**, 4245 (1998).
10. A. Eldar, D. Rosin, B. Z. Shilo, N. Barkai, *Dev. Cell* **5**, 635 (2003).
11. A. D. Lander, Q. Nie, F. Y. Wan, *Bull. Math. Biol.* **69**, 33 (2007).
12. B. Ohkawara, S. Iemura, P. ten Dijke, N. Ueno, *Curr. Biol.* **12**, 205 (2002).
13. T. Bollenbach, K. Kruse, P. Pantazis, M. González-Gaitán, F. Jülicher, *Phys. Rev. Lett.* **94**, 018103 (2005).
14. K. Kruse, P. Pantazis, T. Bollenbach, F. Jülicher, M. González-Gaitán, *Development* **131**, 4843 (2004).
15. A. J. Zhu, M. P. Scott, *Genes Dev.* **18**, 2985 (2004).
16. F. Hsiung, F.-A. Ramirez-Weber, D. D. Iwaki, T. B. Kornberg, *Nature* **437**, 560 (2005).
17. S. R. Yu *et al.*, *Nature* **461**, 533 (2009).
18. G. Schwank *et al.*, *PLoS Biol.* **9**, e1001111 (2011).
19. S. Zhou *et al.*, *Curr. Biol.* **22**, 668 (2012).
20. P. Müller *et al.*, *Science* **336**, 721 (2012); 10.1126/science.1221920.
21. H. Jönsson, M. G. Heisler, B. E. Shapiro, E. M. Meyerowitz, E. Mjolsness, *Proc. Natl. Acad. Sci. U.S.A.* **103**, 1633 (2006).
22. J. Lewis, *Science* **322**, 399 (2008).
23. D. A. Rusakov, D. M. Kullmann, *Proc. Natl. Acad. Sci. U.S.A.* **95**, 8975 (1998).
24. T. F. Schilling, Q. Nie, A. D. Lander, *Curr. Opin. Genet. Dev.* **22**, 562 (2012).
25. R. J. White, Q. Nie, A. D. Lander, T. F. Schilling, *PLoS Biol.* **5**, e304 (2007).
26. A. Linville, K. Radtke, J. S. Waxman, D. Yelon, T. F. Schilling, *Dev. Biol.* **325**, 60 (2009).
27. A. Q. Cai *et al.*, *Development* **139**, 2150 (2012).
28. A. D. Lander, Q. Nie, F. Y. Wan, *Dev. Cell* **2**, 785 (2002).
29. J. Hasty, J. Pradines, M. Dolnik, J. J. Collins, *Proc. Natl. Acad. Sci. U.S.A.* **97**, 2075 (2000).
30. L. Zhang *et al.*, *Mol. Syst. Biol.* **8**, 613 (2012).
31. N. Balaskas *et al.*, *Cell* **148**, 273 (2012).
32. D. M. Umulis, M. Serpe, M. B. O'Connor, H. G. Othmer, *Proc. Natl. Acad. Sci. U.S.A.* **103**, 11613 (2006).
33. A. Goldbeter, D. Gonze, O. Pourquié, *Dev. Dyn.* **236**, 1495 (2007).
34. T. R. Sokolowski, T. Erdmann, P. R. ten Wolde, *PLOS Comput. Biol.* **8**, e1002654 (2012).
35. T. F. Schilling, *Integr. Comp. Biol.* **48**, 658 (2008).
36. M. B. O'Connor, D. Umulis, H. G. Othmer, S. S. Blair, *Development* **133**, 183 (2006).
37. B. Reversade, E. M. De Robertis, *Cell* **123**, 1147 (2005).
38. T. Gregor, D. W. Tank, E. F. Wieschaus, W. Bialek, *Cell* **130**, 153 (2007).
39. M. Zecca, G. Struhl, *PLoS Biol.* **8**, e1000386 (2010).
40. A. Casali, G. Struhl, *Nature* **431**, 76 (2004).
41. P. McHale, W. J. Rappel, H. Levine, *Phys. Biol.* **3**, 107 (2006).
42. D. Rogulja, C. Rauskolb, K. D. Irvine, *Dev. Cell* **15**, 309 (2008).
43. N. A. Monk, *Bull. Math. Biol.* **60**, 901 (1998).
44. F. A. Grusche, H. E. Richardson, K. F. Harvey, *Curr. Biol.* **20**, R574 (2010).
45. S. Kondo, T. Miura, *Science* **329**, 1616 (2010).
46. R. A. Foty, M. S. Steinberg, *Dev. Biol.* **278**, 255 (2005).
47. M. Howard, K. Kruse, *J. Cell Biol.* **168**, 533 (2005).

Acknowledgments: Supported by NIH grants P50GM076516 and R01GM067247.

Supplementary Materials

www.sciencemag.org/cgi/content/full/339/6122/923/DC1
Materials and Methods
References

10.1126/science.1224186

Using the Earth as a Polarized Electron Source to Search for Long-Range Spin-Spin Interactions

Larry Hunter,[1]* Joel Gordon,[1] Stephen Peck,[1] Daniel Ang,[1] Jung-Fu Lin[2]

Many particle-physics models that extend the standard model predict the existence of long-range spin-spin interactions. We propose an approach that uses the Earth as a polarized spin source to investigate these interactions. Using recent deep-Earth geophysics and geochemistry results, we create a comprehensive map of electron polarization within the Earth induced by the geomagnetic field. We examine possible long-range interactions between these spin-polarized geoelectrons and the spin-polarized electrons and nucleons in three laboratory experiments. By combining our model and the results from these experiments, we establish bounds on torsion gravity and possible long-range spin-spin forces associated with the virtual exchange of either spin-one axial bosons or unparticles.

Advances in particle physics have contributed to our understanding of the deep Earth, most recently through the observation of geoneutrinos (*1*). By contrast, the contributions to particle physics resulting from our understanding of the Earth have largely been limited to gravitational interactions or using the Earth as a large mass or baryonic source. Here we suggest that our knowledge of the magnetic fields and electron-spin behavior within the Earth is sufficiently developed that we can use the Earth's polarized electrons to study anomalous long-range spin-spin interactions.

Many extensions of the standard model of particle physics predict the existence of new particles. The virtual exchange of these particles between ordinary fermions can result in spin-spin interactions that look quite different from those expected from electromagnetism. The spin-dependent forces that result from the exchange of a pseudoscalar boson like the axion were originally investigated by Moody and Wilczek (*2*). The exchange of a vector boson with mass $m_{z'}$ can lead to the "spin-dot-spin" and "spin-cross-spin" potentials (*3*)

$$V_1 = \frac{g_A^1 g_A^2}{4\pi r}(\hat{\sigma}_1 \cdot \hat{\sigma}_2)e^{-r/\lambda} \qquad (1)$$

$$V_2 = \frac{\hbar^2}{4\pi}\left(\frac{g_V^1 g_A^2}{2M_1} + \frac{g_A^1 g_V^2}{2M_2}\right)(\hat{\sigma}_1 \times \hat{\sigma}_2)\cdot\hat{r}\left(1+\frac{r}{\lambda}\right) \times \left(\frac{1}{r^2}\right)e^{-r/\lambda} \qquad (2)$$

where g denotes the vector (V) or axial (A) coupling constants of fermions 1 or 2 with mass M and spin directions, $\hat{\sigma}$. The interaction range of the force is denoted by $\lambda = \hbar/m_{z'}c$ where $\hbar$ is Planck's constant (h) divided by 2π and c is the speed of light.

Another interesting entity that could produce long-range spin-spin interactions is the "unparticle" (*4*). Unlike ordinary particles, unparticles do not have well-defined masses but can be characterized in terms of an energy scale Λ, a scaling dimension d, and a dimensionless coupling constant c_A. In the long-range limit, the virtual exchange of an axial-vector unparticle results in an effective potential (*5*)

$$V_u = -c_A^2 \frac{4\sqrt{\pi}\Gamma(d+1/2)\Gamma(2(d-1))}{(2\pi)^{2d}\Gamma(d-1)\Gamma(2d)}(\hat{\sigma}_1 \cdot \hat{\sigma}_2) \times \left(\frac{\hbar c}{r}\right)\left(\frac{\lambda_u}{r}\right)^{2d-2} \qquad (3)$$

where Γ is the gamma function, and $\lambda_u = \hbar c/\Lambda$ is the characteristic length associated with the unparticle. The unparticle potential does not exhibit the usual exponential decay with distance that is associated with the exchange of ordinary massive particles.

Ramsey conducted the first experiment that looked for anomalous spin-spin couplings (*6*). Early work in the field placed limits on such couplings between electrons (e^--e^-) (*7*, *8*) and between electrons and nucleons (*9*). Recently, constraints have been placed on short-range (atomic scale) anomalous spin-spin forces by considering hyperfine structure in hydrogen-like atoms (*10*), spin-exchange collisions (*11*), and magnetic resonance in deuterated molecular hydrogen (*12*). New searches for long-range interactions have been quite successful in placing bounds on the coupling constants associated with the potentials discussed above. Neutron-neutron (n-n) spin couplings have been investigated by using nuclear-spin magnetometers with spin-polarized ^{3}He used as a source (*13*, *14*), whereas the e^--e^- spin interactions have been investigated by using a spin-polarized torsion pendulum with $SmCo_5$-Alnico toroids as sources (*15*). To search for long-range ($\lambda > \sim 1$ m) spin-spin interactions, the polarized source is placed near the apparatus and the direction of its polarization is altered. The spin associated with optically pumped nuclear sources can be reversed simply by changing the circular polarization of the pumping radiation. The solid-state electron-spin sources can be easily rotated in the laboratory. One then looks for a response in the magnetometer or torsion pendulum that is synchronous with the source modulation. Great efforts are made in these experiments to minimize the ordinary magnetic-dipole coupling between the source and apparatus.

The concept of the experiment. Here we suggest an alternative approach where, instead of the modulated laboratory spin source, one uses the polarized electrons within the Earth. The advantage of this approach is simply one of numbers. There are $\sim 10^{49}$ unpaired electron spins in the Earth. On average, about one extra electron out of every 10 million will become polarized antiparallel to Earth's magnetic field. Hence, there are on the order of 10^{42} polarized electrons in the Earth, compared with $\sim 10^{22}$ polarized neutrons or $\sim 10^{25}$ polarized electrons in a typical laboratory source. Hence, the number of polarized geoelectrons exceeds the number in a laboratory source by a factor of at least 10^{17}. Laboratory sources are usually located a few tenths of a meter from the detection apparatus while a typical geoelectron is a few thousand kilometers away. For ordinary electromagnetism and for the spin-spin potentials associated with pseudoscalar bosons, the dipole-dipole interaction falls off as the cube of the distance, r, and the increased distance of the geoelectrons reduces the interaction potential by ~21 orders of magnitude. For such potentials there is no net advantage in considering an Earth source. However, many of the anomalous spin-spin potentials (e.g., Eqs. 1 to 3) fall off as $1/r^n$, where n is between 1 and 2. For these potentials, the suppression of the sensitivity by the distance will be between 7 and 14 orders of magnitude. Here, even with additional losses due to poor geometry and lower experimental sensitivity, a geoelectron source can result in substantially improved constraints.

The greatest disadvantage to searching for spin interactions with the Earth is that one cannot modulate the spin source. To extract the geoelectron spin-spin signal, one must reverse or modulate the contribution it makes to the experimental signal. This can be accomplished by mounting the detection apparatus on a rotating table. Such systems have been developed for searches for violations of local Lorentz invariance (LLI) (*16*, *17*). Although most contributions to the experimental signal are independent

[1]Physics Department, Amherst College, Amherst, MA 01002, USA. [2]Department of Geological Sciences, Jackson School of Geosciences, The University of Texas at Austin, Austin, TX 78712, USA.

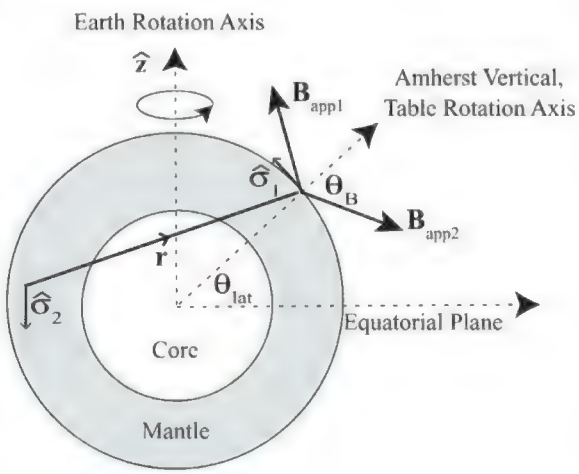
*To whom correspondence should be addressed. E-mail: lrhunter@amherst.edu

of the table orientation, the "effective" spin associated with the geoelectrons creates a preferred axis in the laboratory that couples to the table position via the various spin-spin interactions (Eqs. 1 to 3).

Energy bounds on spins oriented relative to Earth. We use the measurements reported from two recent LLI experiments [one in Amherst, MA (θ_{lat} = 42.37°N, 72.53°W) (*18*) and another in Seattle, WA (47.658°N, 122.3°W) (*19*)], and bounds from an earlier Seattle experiment that searched for the coupling of nuclear spin to Earth's gravitational field (*20*), to extract limits on the associated spin-spin couplings.

The geometry of the Amherst ^{199}Hg-Cs comagnetometer experiment is shown in Fig. 1. The experiment is mounted on a table that rotates between two data-collection positions separated by 180°. The table rotation changes the direction of the applied magnetic field vector from $\mathbf{B}_{app1}$ to $\mathbf{B}_{app2}$. This inverts the horizontal component of the magnetic field while leaving the vertical component fixed. The Cs magnetometer is used to hold the magnitude of the magnetic field constant. The change in the ^{199}Hg nuclear precession frequency with the table position is measured to be $\Delta\nu_N^{Hg} < 1.1$ μHz (2σ bounds will be quoted throughout this paper). The resulting bound on the energy of a ^{199}Hg nuclear spin ($\hat{\sigma}_1$ in Fig. 1) oriented north (N) is $\hat{\beta}_N^{Hg} < h\Delta\nu_N^{Hg}/(4\sin\theta_B) = 1.3 \times 10^{-21}$ eV, where the geometric factor is associated with the magnetic field angle with respect to vertical (θ_B = 63.8°) and the factor of 4 accounts for the differences between the two table positions and the two orientations of the nuclear spin with respect to the applied field. We follow the notation of (*19*) in using a "hat" over β to indicate that a correction for Earth's gyroscopic frequency has been applied. A random geoelectron with a spin $\hat{\sigma}_2$, separated from the apparatus by a distance *r*, is also shown in the plot. To place bounds on the coupling strengths in Eqs. 1 and 3, we sum the interaction potentials over all of Earth's spin-polarized electrons and require that the resulting energy not exceed $\hat{\beta}_N^{Hg}$.

Fig. 1. The geometry of the Amherst experiment (see text for explanation).



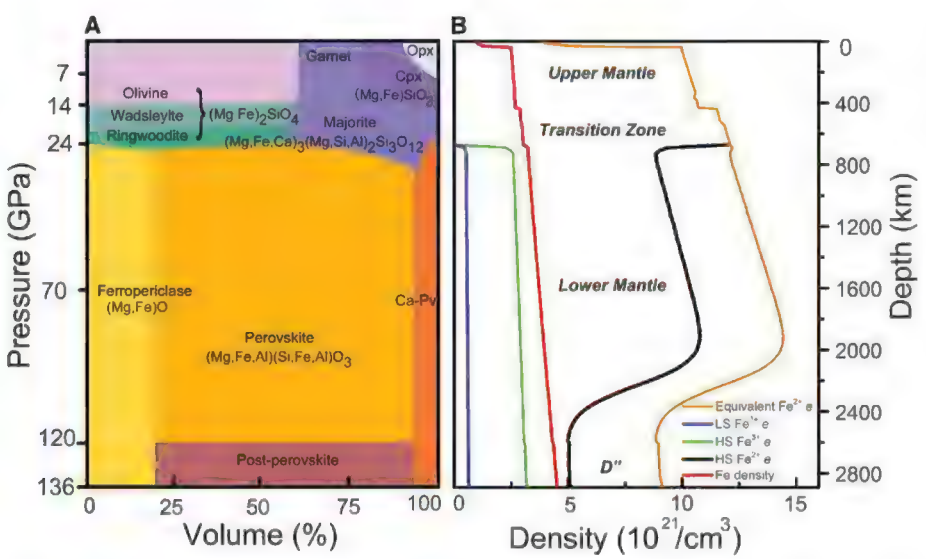


Fig. 2. (**A**) Mineral proportions in the pyrolite model as a function of depth and (**B**) the resulting iron and electron densities in the various electronic spin states. The PREM model pressure profile is used throughout the calculations (*30*). The volume fractions and the iron fractions of the mantle are from (*23*), while the densities and iron fractions for the crust are from (*28*). The "Equivalent Fe^{2+} e^-" density combines the electrons in the different spin states, weighted by a factor that takes into account the relevant proportion of their contribution to the polarized electron-spin density. For the *D"* layer with the silicate post-perovskite phase between the core and the mantle, we use the temperature gradient suggested in (*29*).

The coupling of the geoelectron spins by the spin-cross-spin potential (Eq. 2) is about an order of magnitude larger for detection spins oriented east (E) than for those oriented north. To improve our bounds on this potential, we have made additional measurements with our apparatus (*18*) rotated 90°. In this orientation, the horizontal component of the applied magnetic field changes from east to west when the table changes position. We measure the change in the ^{199}Hg nuclear precession frequency with the table position to be $\Delta\nu_E^{Hg} < 0.8$ μHz. The resulting bound on the energy of a ^{199}Hg nuclear spin oriented east is $\hat{\beta}_E^{Hg} < h\Delta\nu_E^{Hg}/(4 \times \sin(63.8°)) = 9 \times 10^{-22}$ eV. We use this bound to constrain the strength of the spin-cross-spin potential.

In their search for a cosmic preferred frame, a recent Seattle experiment constrains possible directional couplings to the polarized electrons that make up their $SmCo_5$-Alnico torsion pendulum. Their bounds on a spin potential coupling electron spin along the north and east directions are $\hat{\beta}_N < 5.9 \times 10^{-21}$ eV and $\hat{\beta}_E < 8 \times 10^{-22}$ eV. Both of these bounds can be used to set limits on electron-electron spin-spin couplings of Eqs. 1 to 3. In practice, the superior experimental bound along the east direction yields the best electron limits on all three potentials.

The third experiment that we have considered was explicitly designed to search for a coupling between local gravity and nuclear spin. The experiment compares the precession frequencies of ^{199}Hg and ^{201}Hg in a comagnetometer arrangement. This experiment is not mounted on a rotating table, but instead modulates the detection by periodically reversing the direction of the applied magnetic field and occasionally inverting the orientation of the apparatus. Unlike the LLI experiments, where the sensitive directions lie in the horizontal plane, the magnetometer field is oriented along Earth's spin axis (*z*). Using an updated nuclear structure calculation (*21*), the experiment places a bound on the neutron (*n*) and proton (*p*) spin couplings along the *z* axis, $\hat{B}_z^n < \frac{1}{2}(3.6 \times 10^{-20}\ \text{eV})\cos(47.5°) = 1.2 \times 10^{-20}$ eV and $\hat{B}_z^p < \frac{1}{2}(5.2 \times 10^{-20}\ \text{eV})\cos(47.5°) = 1.8 \times 10^{-20}$ eV. Here the geometrical factor transforms the bound from along the vertical to along *z* and the factor of 2 accounts for the two orientations of the nuclear spins with respect to the applied field.

Electron-spin density map. To use Earth as a spin source requires creating a map of the polarized electron-spin density and direction everywhere within the Earth. Three input parameters are needed to create this map: temperature, direction and magnitude of Earth's magnetic field, and the density of unpaired electrons that can be aligned in the external field. We consider each of these in the following discussion.

Earth's core is mostly made of metallic Fe with ~5% Ni, together with some lighter elements. First-principles calculations indicate that there is no paramagnetism associated with the Fe-Ni alloy in the core (*22*). Hence, the core makes no contribution to Earth's polarized spin density and will be neglected in our model.

The electron-spin density in Earth's mantle can be calculated by using recent results from deep-Earth mineral physics, geochemistry, and seismology. The volume fractions and iron contents of various regions of the mantle have been reported experimentally (Fig. 2) (*23*). Because we are interested in establishing upper bounds on the spin-spin coupling constants, we shall make a conservative estimate of the magnitude of the electronic spins within the Earth. Iron, the most abundant transition metal in oxides and silicates of Earth's mantle and crust, has a partially filled d-shell that dominates the resulting paramagnetism of the mantle minerals and crustal rocks. Other major rock-forming elements such as Mg, Si, Al, and O with closed electron shells have negligible contributions to the polarized spin density. We use a representative pyrolite compositional model of the mantle that assumes there is no major chemical difference between the elemental composition of the upper and lower mantle (*24*). Alternative models generally yield higher iron concentrations in the lower mantle (*25*). We use the Birch-Murnaghan equation of state (*26*), relevant physical parameters of the mantle silicates and oxides, and expected pressure-temperature profiles of the crust and mantle (geotherm) to calculate the densities of the major mantle minerals at different depths (see table S1 for details) (*27*–*29*). In fig. S1, we compare our calculated mean density of these minerals at various depths with the seismic preliminary reference Earth model (PREM) values (*30*).

Most of the iron in Earth's crust and mantle is bound up in mineral lattices with either ferrous (Fe^{2+}) or ferric (Fe^{3+}) valence states. For the crust and upper mantle, most of the iron exists in the Fe^{2+} state with four unpaired d-shell electrons and a total spin $S = 2$, the so-called high-spin (HS) state. The lower mantle has two primary iron-bearing constituents, ferropericlase (fp) and perovskite (pv), whereas silicate post-perovskite (ppv) likely exists in the lowermost mantle (the D″ region) (Fig. 2). Ferropericlase consists predominantly of ferrous iron while pv and ppv consist of approximately half Fe^{2+} and half Fe^{3+}. In the lower mantle, fp undergoes a transition from its HS state to a low-spin (LS) state with $S = 0$ (*31*). Along an expected lower-mantle geotherm (*32*), experimental results show that the spin crossover of Fe^{2+} in fp starts at ~70 GPa (1700-km depth) and 2200 K and completes at ~125 GPa (2700-km depth) and 2400 K (*33*). This spin crossover reduces the unpaired electron density at depths greater than ~2100 km (Fig. 2). The spin states of iron in pv and ppv remain under debate, but the current consensus is that a similar spin transition from $S = 5/2$ to $S = 1/2$ is expected to occur in Fe^{3+} in the octahedral site (B site) of pv and ppv whereas Fe^{2+} remains in the HS state at all mantle pressures (*34*). We choose again to conservatively assume that all of the Fe^{3+} in the B site of pv and ppv, which is ~25% of all iron in those minerals, is in the LS state, whereas the remainder of the Fe^{3+} and Fe^{2+} remains in the HS states. The resulting unpaired electron densities as a function of depth are shown in Fig. 2.

For our model to be successful, the magnitude and direction of Earth's magnetic field, **B**, must be known everywhere between the core and the surface. The magnetic field within the Earth is predominantly created by electrical currents within the core (*35*). We use the world magnetic model (WMM 2010) (*36*) to recreate the magnetic field within the Earth as it existed when the data were collected. The magnetic field is derived from a magnetic potential that is approximated by a 12th-order associated Legendre expansion. The coefficients of the expansion are chosen to yield a best fit to worldwide and satellite observations. The model is expected to be valid throughout the mantle and crust, where the material is believed to be electrically insulating. For the Amherst data, we model the field as it was on 1 October 2010, while for the Seattle data we use 1 July 2007 for the torsion pendulum experiment and 1 January 1991 for the Hg comagnetometer experiment. In practice, Earth's magnetic variation is sufficiently slow that altering the chosen date by a decade does not produce any notable change in our results.

When an iron ion in the mantle interacts with this external field, the spin becomes slightly polarized. We assume that the orbital moments are quenched and can be neglected. The average value of the spin of each HS Fe^{2+} (in units of $\hbar$) along the magnetic field is $\langle S_z \rangle_{Iron} = 2g\mu_B B/k_B T$, where μ_B is the Bohr magneton, k_B is Boltzmann's constant, T is the temperature, and g is the electron g factor. Because on average, each of the four unpaired electrons has a projection of its spin along z that is ¼ of that for the lattice site, the effective alignment of each electron in the

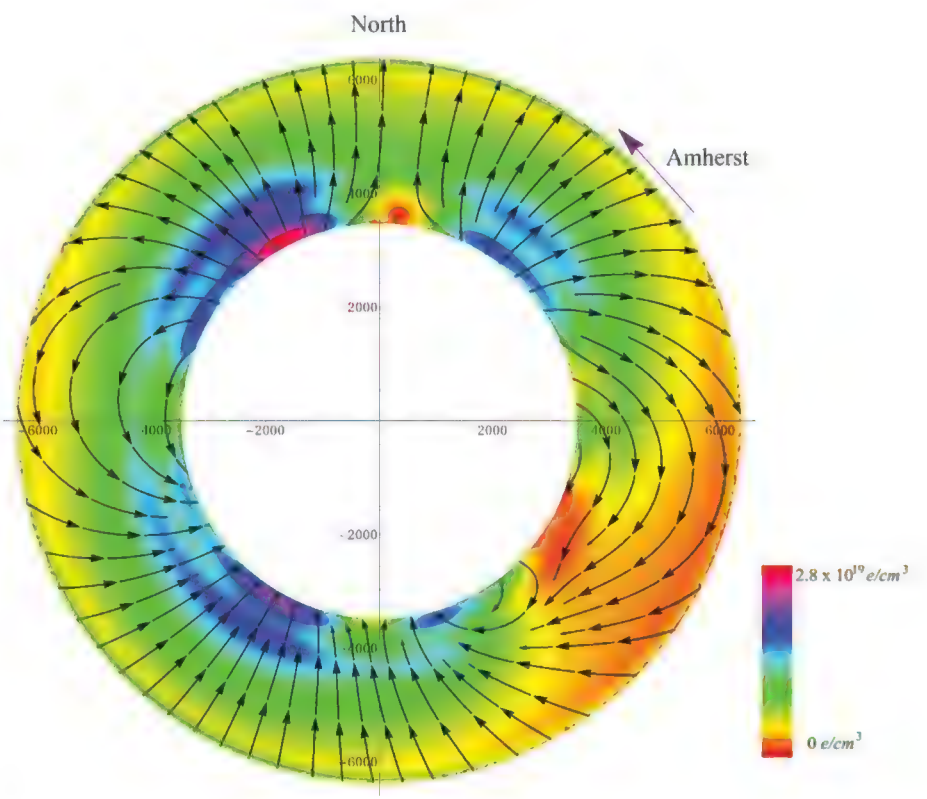


Fig. 3. The polarized electron-spin density on a plane that contains Earth's rotation axis and Amherst, MA. The black arrows indicate only the direction of the electron spins, while the color shading of the plot indicates the magnitude of the polarized electron-spin density. The density of the arrows on the plot is not meaningful. The $\hat{\phi}$ (east) component of the field (into or out of the page) is not shown. The electron-spin density within the core (white central circle) is assumed to be zero. The violet vector corresponds to the north orientation and location of the Amherst apparatus. The vertical z axis is along Earth's rotation axis, and the axis labels have units of kilometers.

external field is $\langle S_z \rangle_e = g\mu_B B/2k_B T$. For a spin ½ electron in HS Fe^{2+}, this corresponds to an average spin alignment of $2\mu_B B/k_B T$, where we assume $g = 2$. Similar considerations for the LS ($S = 1/2$) and HS ($S = 5/2$) states of Fe^{3+} yield average electron spin alignments of $\mu_B B/k_B T$ and $7\mu_B B/3k_B T$, respectively (*37*). Figure 1 shows the "equivalent Fe^{2+} e^-" density, where we have taken this weighting into account. The polarized electron-spin density is determined by simply multiplying this density by $2\mu_B B/k_B T$. For a field of 1 G and a temperature of 2000 K, a typical HS Fe^{2+} electron's polarization is about 7×10^{-8}.

Combining the magnetic field map with the temperature and electron-spin density profiles yields a map of the polarized electron-spin density and direction within the Earth. A sampling of the result on a plane that includes Earth's rotation axis and Amherst is shown in Fig. 3. The general rise in the magnitude of the polarization with decreasing radius is due to the increased magnitude of the magnetic field and the higher density of the constituent minerals. When moving in toward the core, the abrupt drop in the polarized spin density at a radius of about 4200 km is associated with the spin transition in fp, whereas the drop just before reaching the core is a result of the rising temperature. The slight rise in the spin density at the crust is associated with the rapidly decreasing temperature as one approaches the surface. The reduction of the spin density in the fourth quadrant of the plot results from a minimum in the Earth field commonly referred to as the South Atlantic Anomaly (*38*). Also shown on the plot is a vector corresponding to the north orientation of the Amherst apparatus ($\hat{\sigma}_1$ in Fig. 1). In this orientation, only geoelectron spin components parallel to this direction contribute to the dot product between the source and detection spins (Eqs. 1 and 3). Cross-product contributions (Eq. 2) are maximized when the detector orientation is east, which is into the page at Amherst on Fig. 3.

Interaction potentials. With the polarized geoelectron spin density and direction known everywhere within the Earth, we can calculate the interaction potential associated with each of the possible spin interactions described in Eqs. 1 to 3. The various integrations over the Earth volume are run in Mathematica using geocentric coordinates (see fig. S2 for details). For $1/r$ and $1/r^2$ potentials, the relative contributions to the integral from various depths are shown in fig. S3. For the $1/r$ ($1/r^2$) potential, the relative contributions to the integral from the crust, upper mantle, and lower mantle are 2, 31, and 66% (6, 43, and 51%), respectively. To limit couplings between the geoelectrons and ^{199}Hg nuclei, we require that the integrated spin-spin potentials (relative to Amherst) remain smaller than the observed experimental upper bound: $\hat{\beta}_N^{Hg}$ for the spin-dot-spin potentials (Eqs. 1 and 3) and $\hat{\beta}_E^{Hg}$ for the spin-cross-spin potential (Eq. 2). To extract the bounds on the electron coupling to the neutron and proton, we assume that the ^{199}Hg nuclear spin has a neutron spin projection of –31% and a proton spin projection of –3% (*21*). The bounds on the electron-electron couplings and additional bounds on the electron-neutron and electron-proton couplings are similarly derived from integrations relative to Seattle and the bounds β_E^e, $\hat{\beta}_z^n$, and $\hat{\beta}_z^p$. The results for a potential mediated by an intermediate vector boson of mass $m_{z'}$ are shown in Fig. 4. Limits on the unparticle couplings are shown in Table 1.

Our predicted bounds are markedly insensitive to changes in various Earth-model parameters. The bounds in Fig. 4 are proportional to the absolute temperature and inversely proportional to the densities assumed. Hence, a 20% change in either the density or the temperature would change the bounds by about 20%. Such large departures from the pyrolite and PREM values used here seem highly unlikely. Indeed, Amherst should be reasonably well represented by the average continental values assumed. Seattle resides on top of a deep, ancient subducting plate that may produce appreciable local cooling (*39*). However, lower temperatures would only further improve the bounds obtained from the model. We do not report any limits from our model for ranges less than 1 km where local density variations and the possibilities of local ferromagnetic materials render the results unreliable. On the logarithmic plots of Fig. 4, a 20% change would result in a shift that is less than the thickness of the plotted lines.

Discussion. Our results establish limits on many of the possible long-range spin-spin interactions between fermions. For the axial-axial exchange (Eq. 1), the low-mass limits achieved are about a million times lower than the *n-n* limits that were established with a laboratory source. In the long-range limit, the potential has the same $1/r$ dependence as gravity, and one can calculate the ratio between the strength of the spin-spin coupling between two particles to their gravitational

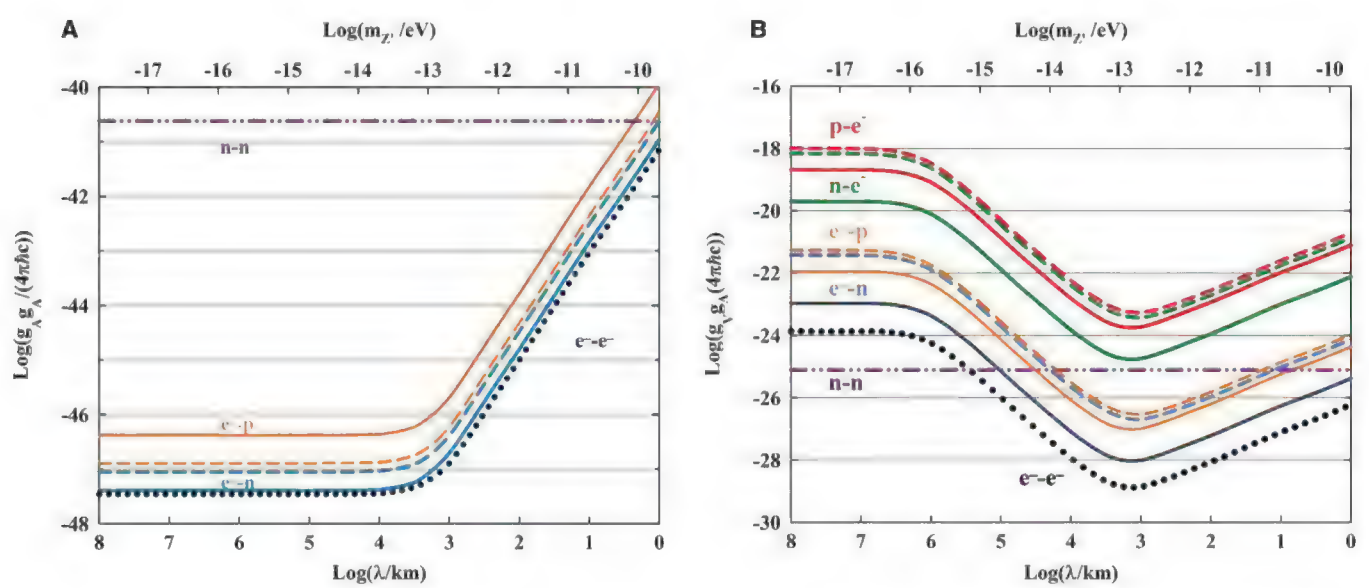


Fig. 4. (**A**) Bounds on the long-range axial-axial couplings (Eq. 1) as a function of the range and mass of the intermediate vector boson. All regions of the plots above the lines are excluded. The violet lines (mixed dots and dashes) are the *n-n* bound established by (*13*). All other bounds are from the present analysis. We extract the e^--e^- bounds (black dots) from the data in (*19*) and the e^--*p* (orange) and e^--*n* (blue) bounds from the data in (*18*) (solid lines) and (*20*) (dashed lines). (**B**) Bounds on the vector (V)–axial (A) couplings (Eq. 2). The first fermion of each labeled pair has the vector coupling, whereas the second has the axial coupling. The violet and black limits are derived as discussed in (A). We extract the e^--*p* (orange), e^--*n* (blue), *n*-e^- (green), and *p*-e^- (brown) bounds from our new measurement of β_E^{Hg} (solid lines) and the data in (*20*) (dashed lines). Here e^-, *n*, and *p* correspond to the electron, neutron, and proton.

Table 1. Upper bounds on the axial dimensionless coupling c_A associated with unparticles with scaling dimension d. Here e^-, n, and p denote the couplings between electrons, neutrons, and protons. The n-n limit was established in (*13*). All other bounds are from the present analysis. We extract the e^--n, e^--p, and e^--e^- bounds from the data in (*18*), (*20*), and (*19*), respectively. For ease of comparison with the n-n results, we have set $\Lambda = 1$ TeV in this analysis.

c_A|d	1	1.125	1.25	1.33	1.375	1.5
e^--n	7×10^{-24}	1×10^{-20}	2×10^{-17}	2×10^{-15}	3×10^{-14}	6×10^{-11}
e^--p	1×10^{-23}	2×10^{-20}	4×10^{-17}	4×10^{-15}	6×10^{-14}	1×10^{-10}
e^--e^-	7×10^{-24}	1×10^{-20}	2×10^{-17}	2×10^{-15}	3×10^{-14}	5×10^{-11}
n-n	1×10^{-20}		1×10^{-15}	4×10^{-14}		8×10^{-11}

attraction. Our bounds imply that this ratio is <0.2% for e^--e^- interactions and $<1.2 \times 10^{-6}$ for e^--n interactions. Quantum gravity can induce an effective four-fermion axial interaction with a characteristic coupling of $M_1 M_2 M_p^2$ where M_1 and M_2 are the fermion masses and $M_p \approx 1.2 \times 10^{-19}$ GeV/c^2 is the Planck mass (*40*–*42*). For two electrons, this yields a dimensionless coupling of 10^{-45}, whereas for a neutron and an electron the characteristic coupling is 10^{-42}. Our measured long-range bounds on these couplings (Fig. 3) are less than 10^{-47}, suggesting that our results could be sensitive to such quantum gravity effects.

For the vector-axial coupling of the spin-cross-spin potential (Eq. 2), the range dependence of the limits is more complex. Our limits improve on the bounds obtained by laboratory sources (by as much as a factor of 4000) but only for ranges, ~1 km < λ < ~200,000 km. The integration exhibits a high degree of cancellation for ranges significantly larger than Earth's diameter, suggesting a contribution from Earth's more distant hemisphere of opposite sign and nearly equal magnitude to that from the near hemisphere.

For the unparticle exchange with $d = 1$, the potential falls off as $1/r$, and the bounds on the e^--n and e^--e^- coupling amplitudes are more than a thousand times more restrictive than all previous results. Unparticles with $d = 1.5$ correspond to a dot product potential that falls off as $1/r^2$. Here our bounds become similar to those that have been achieved with laboratory sources.

Outlook. Because the spins in our apparatus on Earth's surface have a higher eastward velocity than most of the spins within the mantle, the geoelectron source offers the possibility of testing all of the velocity-dependent effective potentials proposed in (*3*). Experiments using laboratory sources are not sensitive to these terms, and most of these possible spin-spin potentials are currently unbounded. Ongoing development of comagnetometers and nuclear gyroscopes (*43*) should result in improved experimental sensitivities that will allow further refinement of the electron-nucleon couplings. With a different set of assumptions, it may be possible to extend this model so that Earth might be used as a source of nuclear spin. This might allow more stringent limits to be established on nucleon-nucleon spin couplings. Performing the experiment in a region where the surface magnetic field is stronger and more parallel to the surface can substantially enhance sensitivity. An identical experiment conducted in southern Thailand would have about twice the sensitivity of one conducted in Amherst.

If refined experiments find evidence for a long-range spin-spin interaction, the discovery could provide a powerful tool for the investigation of the mantle. Once the potential is well understood, measurements similar to those discussed here could be used to probe Earth's iron concentration and its spin and valence states as a function of depth (*31*, *33*). Such information might eventually help reconcile seismic observations and mineral physics data with geochemical models (*23*, *25*, *44*).

References and Notes

1. A. Gando *et al.*, *Nat. Geosci.* **4**, 647 (2011).
2. J. E. Moody, F. Wilczek, *Phys. Rev. D Part. Fields* **30**, 130 (1984).
3. B. A. Dobrescu, I. Mocioiu, *J. High Energy Phys.* **11**, 005 (2006).
4. H. Georgi, *Phys. Rev. Lett.* **98**, 221601 (2007).
5. Y. Liao, J.-Y. Liu, *Phys. Rev. Lett.* **99**, 191804 (2007).
6. N. F. Ramsey, *Physica* **96A**, 285 (1979).
7. T. C. P. Chui, W.-T. Ni, *Phys. Rev. Lett.* **71**, 3247 (1993).
8. V. F. Bobrakov *et al.*, *JETP Lett.* **53**, 294 (1991).
9. D. J. Wineland, J. J. Bollinger, D. J. Heinzen, W. M. Itano, M. G. Raizen, *Phys. Rev. Lett.* **67**, 1735 (1991).
10. S. G. Karshenboim, *Phys. Rev. A* **83**, 062119 (2011).
11. D. F. Jackson Kimball, A. Boyd, D. Budker, *Phys. Rev. A* **82**, 062714 (2010).
12. M. P. Ledbetter, M. V. Romalis, D. F. Jackson Kimball, Constraints on short-range spin-dependent interactions from scalar spin-spin coupling in deuterated molecular hydrogen. *Phys. Rev. Lett.* **110**, 040402 (2013).
13. G. Vasilakis, J. M. Brown, T. W. Kornack, M. V. Romalis, *Phys. Rev. Lett.* **103**, 261801 (2009).
14. A. G. Glenday, C. E. Cramer, D. F. Phillips, R. L. Walsworth, *Phys. Rev. Lett.* **101**, 261801 (2008).
15. C. E. Cramer, A torsion balance search for spin-coupled forces, thesis, Univ. of Washington (2007).
16. L. S. Hou, W. T. Ni, Y. C. M. Li, *Phys. Rev. Lett.* **90**, 201101 (2003).
17. J. M. Brown, S. J. Smullin, T. W. Kornack, M. V. Romalis, *Phys. Rev. Lett.* **105**, 151604 (2010).
18. S. K. Peck *et al.*, *Phys. Rev. A* **86**, 012109 (2012).
19. B. R. Heckel *et al.*, *Phys. Rev. D Part. Fields Gravit. Cosmol.* **78**, 092006 (2008).
20. B. J. Venema, P. K. Majumder, S. K. Lamoreaux, B. R. Heckel, E. N. Fortson, *Phys. Rev. Lett.* **68**, 135 (1992).
21. V. Flambaum, S. Lambert, M. Pospelov, *Phys. Rev. D Part. Fields Gravit. Cosmol.* **80**, 105021 (2009).
22. D. Alfè, M. J. Gillan, *Phys. Rev. B* **58**, 8248 (1998).
23. T. Irifune *et al.*, *Science* **327**, 193 (2010).
24. A. E. Ringwood, *Geochim. Cosmochim. Acta* **55**, 2083 (1991).
25. R. Jeanloz, E. Knittle, *Philos. Trans. R. Soc. Lond. A* **328**, 377 (1989).
26. F. Birch, *J. Geophys. Res.* **91**, (B5), 4949 (1986).
27. J. P. Poirier, *Introduction to the Physics of the Earth's Interior* (Cambridge Univ. Press, Cambridge, ed. 2, 2000).
28. F. D. Stacey, P. M. Davis, *Physics of the Earth* (Cambridge Univ. Press, Cambridge, ed. 4, 2008).
29. T. Lay, J. Hernlund, B. A. Buffett, *Nat. Geosci.* **1**, 25 (2008).
30. A. M. Dziewonski, D. L. Anderson, *Phys. Earth Planet. Inter.* **25**, 297 (1981).
31. J. F. Lin, T. Tsuchiya, *Phys. Earth Planet. Inter.* **170**, 248 (2008).
32. J. M. Brown, T. J. Shankland, *Geophys. J. R. Astron. Soc.* **66**, 579 (1981).
33. Z. Mao, J. F. Lin, J. Liu, V. B. Prakapenka, *Geophys. Res. Lett.* **38**, L23308 (2011).
34. Y. G. Yu, H. Hsu, M. Cococcioni, R. M. Wentzcovitch, *Earth Planet. Sci. Lett.* **331–332**, 1 (2012).
35. B. A. Buffett, *Science* **288**, 2007 (2000).
36. S. Maus *et al.*, "The US/UK World Magnetic Model for 2010-2015" (National Oceanic and Atmospheric Administration, Technical Report NESDros. Inf. Serv./NGDC 2010).
37. C. Kittel, *Introduction to Solid State Physics* (Wiley, ed. 5, 1976).
38. G. Hulot, N. Olsen, T. J. Sabaka, in *Treatise Geophys.*, G. Schubert, Ed. (Elsevier, Amsterdam, 2007), vol. 5, chap. 33.
39. Y. Tian, D. Zhao, *Phys. Earth Planet. Inter.* **200–201**, 72 (2012).
40. D. E. Neville, *Phys. Rev. D Part. Fields* **25**, 573 (1982).
41. T. W. B. Kibble, *J. Math. Phys.* **2**, 212 (1961).
42. I. B. Khriplovich, *Phys. Lett. B* **709**, 111 (2012).
43. E. A. Donley, J. Kitching, Nuclear Magnetic Resonance Gyroscopes in *Optical Magnetometry*, D. Budker, D. F. Jackson-Kimball, Eds. (Cambridge Univ. Press, Cambridge, 2012).
44. M. Murakami, Y. Ohishi, N. Hirao, K. Hirose, *Nature* **485**, 90 (2012).

Acknowledgments: We thank S. Maus, R. Jeanloz, B. Buffett, W. Loinaz, and P. Crowley for useful conversations. We thank A. Anderson, D. Hanneke, and N. Bern for computational and graphics assistance, and C. Lu and J. Yang for compiling mineral physics data and helping with Fig. 2, respectively. We thank N. Fortson and B. Heckel for providing the updated neutron and proton bounds from (*20*). We thank the editors and reviewers for constructive suggestions. This work was supported by funds from the Amherst College Dean of Faculty and the NSF under grants PHY-0855465 and PHY-1205824. J.-F.L acknowledges support from the U.S. NSF (EAR-1056670 and EAR-1053446), Energy Frontier Research in Extreme Environments (EFree), and the Carnegie/DOE Alliance Center (CDAC). The data presented here are available on request from lrhunter@amherst.edu.

Supplementary Materials

www.sciencemag.org/cgi/content/full/339/6122/928/DC1
Figs. S1 to S3
Table S1
References (*45*–*50*)

16 July 2012; accepted 26 December 2012
10.1126/science.1227460

Ferromagnetic Quantum Critical Point in the Heavy-Fermion Metal $YbNi_4(P_{1-x}As_x)_2$

Alexander Steppke,[1*] Robert Küchler,[1] Stefan Lausberg,[1] Edit Lengyel,[1] Lucia Steinke,[1] Robert Borth,[1] Thomas Lühmann,[1] Cornelius Krellner,[1,2] Michael Nicklas,[1] Christoph Geibel,[1] Frank Steglich,[1] Manuel Brando[1*]

Unconventional superconductivity and other previously unknown phases of matter exist in the vicinity of a quantum critical point (QCP): a continuous phase change of matter at absolute zero. Intensive theoretical and experimental investigations on itinerant systems have shown that metallic ferromagnets tend to develop via either a first-order phase transition or through the formation of intermediate superconducting or inhomogeneous magnetic phases. Here, through precision low-temperature measurements, we show that the Grüneisen ratio of the heavy fermion metallic ferromagnet $YbNi_4(P_{0.92}As_{0.08})_2$ diverges upon cooling to $T = 0$, indicating a ferromagnetic QCP. Our observation that this kind of instability, which is forbidden in *d*-electron metals, occurs in a heavy fermion system will have a large impact on the studies of quantum critical materials.

Fluctuations of quantum fields persist to absolute zero. In matter, when such quantum fluctuations are tuned by a nonthermal control parameter such as pressure, they may become sufficiently large to cause a quantum phase transition into a new ground state. If the transition is continuous, it gives rise to a quantum critical point (QCP) (*1*–*3*). QCPs have been observed in a number of antiferromagnetic (AFM) *d*- and *f*-electron metallic systems (*4*, *5*) but are thought not to exist in ferromagnetic (FM) metals for several reasons.

Theoretical considerations for itinerant (such as *d*-electron) systems exclude a FM QCP at field $H = 0$ in clean materials and rather point to a FM quantum phase transition of first order (*6*–*9*). Experiments indicate that FM QCPs are preempted by the development of superconducting (*10*, *11*), modulated (*12*) or spin-glass-like (*13*) phases, or a first-order FM quantum phase transition (*14*). In *f*-electron systems, Kondo screening competes with magnetism and permits QCPs, at which the Kondo effect may even be destroyed, as shown in the AFM heavy-fermion metal $YbRh_2Si_2$ (*15*–*17*). For a FM *f*-derived Kondo-lattice system, it has been shown theoretically that the Kondo screening may completely disappear deep inside the ferromagnetically ordered phase (*18*). It remains to be shown, however, whether the FM QCP concurs also with a Kondo-destroying one.

An ideal method to probe the existence of a heavy fermion QCP (which is generically sensitive to pressure) is to perform combined measurements of the volume thermal expansion, $\beta(T)$, and specific heat, $C(T)$. Both $\beta(T)/T$ and $C(T)/T$ diverge upon approaching the QCP, but the divergence of $\beta(T)/T$ is stronger, leading to a divergence of the Grüneisen ratio $\Gamma(T) = \beta(T)/C(T)$. If the divergence of the Grüneisen ratio follows a power law, $\Gamma(T) \propto T^{-\lambda}$, the critical exponent yields insight into the nature of the QCP (*19*). For instance, in $CeNi_2Ge_2$ $\lambda = 1$ is observed, indicating a three-dimensional (3D) spin-density-wave QCP (*20*). For the unconventional quantum critical material $YbRh_2Si_2$, a fractional exponent, $\lambda \approx 0.7$, was found (*20*).

In this report, we show that in single crystals of the weak 4*f*-derived metallic ferromagnet $YbNi_4P_2$, a FM QCP can be induced by applying negative chemical pressure—substituting 10% As for P. At this QCP, $\beta(T)/T$, $C(T)/T$, and $\Gamma(T)$ diverge with unusual power-law exponents.

Stoichiometric FM metals with a Curie temperature T_C below 1 K are very rare. The Kondo-lattice system $YbNi_4P_2$ is one, with a remarkably low $T_C = 0.17$ K (*21*). This is mainly a result of substantial Kondo screening (with a Kondo temperature $T_K \approx 8$ K) of the magnetic Yb^{3+} ions [Ni is not magnetic (*22*)], leaving a FM ground state with a small ordered moment of ~0.05 Bohr magneton (μ_B) (*23*) and strongly enhanced electron-electron correlations (*21*). It is, therefore, a particularly suitable system to search for and explore FM quantum criticality.

A preliminary study on $YbNi_4P_2$ polycrystals, in which the low-temperature transition at T_C was discovered (*21*), indicated peculiar non-Fermi liquid (NFL) properties in transport and thermodynamic quantities above T_C: in particular, a stronger-than-logarithmically diverging Sommerfeld coefficient $C(T)/T$ and a linear *T*-dependence of the electrical resistivity $\rho(T)$. These results suggest the proximity of a FM QCP; T_C could be tuned smoothly to zero through the isoelectronic substitution of larger As for P, expanding the crystal lattice and causing, in Yb systems, an increase of T_K and a reduction of T_C (*24*). We report thermodynamic and transport measurements on single crystals of stoichiometric $YbNi_4P_2$ (*23*) as well as $YbNi_4(P_{1-x}As_x)_2$, with *x* between 0.04 and 0.13. $YbNi_4P_2$ has a crystalline structure of the tetragonal $ZrFe_4Si_2$ structure type ($P4_2/mnm$), with the Yb^{3+} ions located between chains of edge-connected Ni tetrahedra. The Yb ions form chains along the crystallographic *c* direction, with the lattice parameter $a = 7.0565$ Å being almost twice as large as $c = 3.5877$ Å (Fig. 1A, inset) (*22*). Band-structure calculations show that three sheets of the Fermi surface have

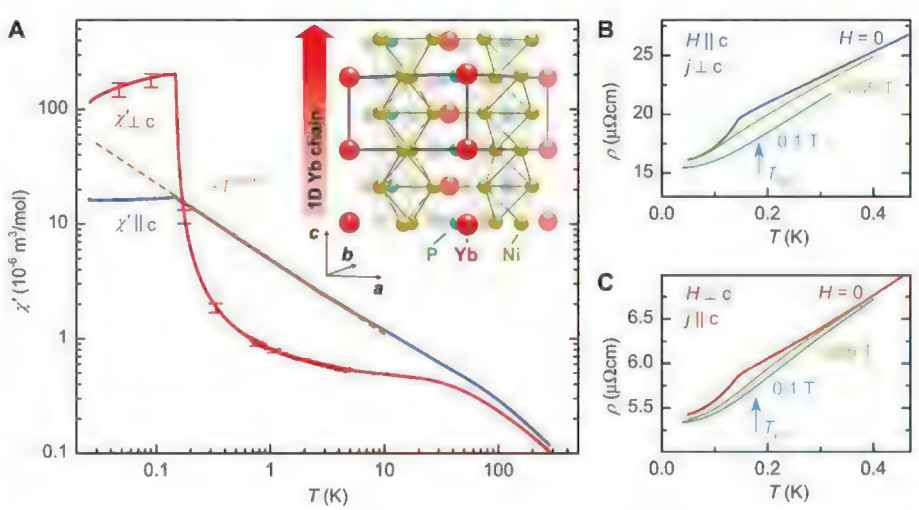


Fig. 1. Temperature dependence of the magnetic and electrical properties of $YbNi_4P_2$. (**A**) Real part of the magnetic susceptibility χ' plotted as a function of temperature, measured perpendicular ($\chi'_\perp$) and parallel ($\chi'_\parallel$) to the *c* axis. The orange dashed line emphasizes the $T^{-0.66}$ behavior of $\chi'_\parallel$ versus *T* below 10 K. (Inset) Unit cell of $YbNi_4P_2$. (**B** and **C**) Electrical resistivity ρ plotted as a function of temperature and measured with current *j* (B) perpendicular and (C) parallel to the crystallographic *c* axis. The kinks at T_C disappear at fields $\mu_0 H \geq 0.05$ T in both field directions. At a field of 0.1 T, both resistivities follow a temperature dependence close to T^2 below T_{FL} (arrows).

[1]Max Planck Institute for Chemical Physics of Solids, Nöthnitzer Strasse 40, 01187 Dresden, Germany. [2]Institute of Physics, Goethe University Frankfurt, Max-von-Laue-Strasse 1, 60438 Frankfurt am Main, Germany.

*To whom correspondence should be addressed. E-mail: steppke@cpfs.mpg.de (A.S.); brando@cpfs.mpg.de (M.B.)

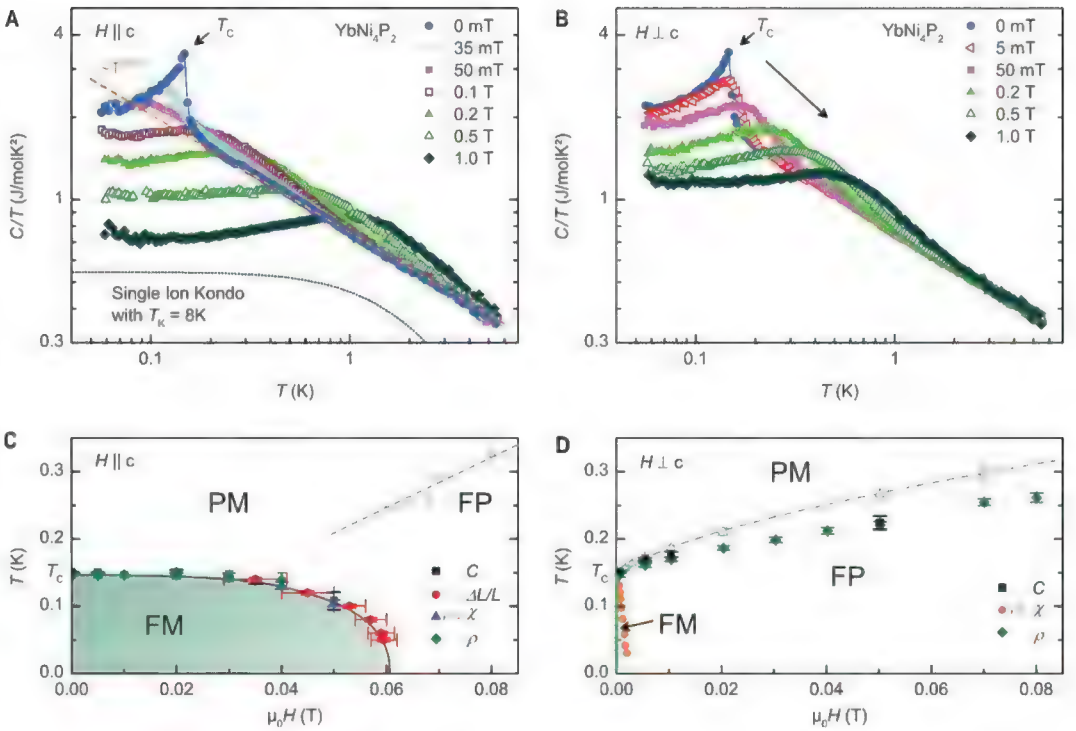


Fig. 2. Magnetic phase diagrams of $YbNi_4P_2$. (**A**) C/T versus T of $YbNi_4P_2$ measured at magnetic fields $H \parallel c$. The dashed orange line emphasizes the $T^{-0.43}$ power-law behavior observed above 0.2 K up to about 6 K. The dotted black line is the specific-heat contribution calculated within the single-ion $S = 1/2$ Kondo model of (*25*), with $T_K = 8$ K and C/T scaled by a factor of 0.5 for better visibility. (**B**) C/T versus T of $YbNi_4P_2$ measured at magnetic fields $H \perp c$. In (A) and (B), the nuclear Schottky contribution has been subtracted from the raw specific-heat data (*32*). (**C**) Magnetic phase diagram for $H \parallel c$, derived from equal-entropy construction for the phase-transition anomaly in the specific-heat coefficient $[C(T)/T]$, kink in the linear magnetostriction $\Delta L(T)/L$ (fig. S10), inflection point of the ac-susceptibility, and first derivative of the resistivity $[\rho(T)]$. The paramagnetic (PM) region is separated from the ferromagnetic (FM) region, where the moments are aligned within the (a,b) plane, by a phase transition of the second order (solid black line). The dashed lines in (C) and (D) indicate the crossover between the PM and the ferromagnetically polarized (FP) regions, determined by the maxima in the ac-susceptibility (fig. S8). (**D**) Magnetic phase diagram for $H \perp c$ derived from $C(T)$, $\rho(T)$, and $\chi'(T)$. The vertical green line at $H = 0$ indicates that the FM order exists at $H = 0$ only. The orange circles mark the increase of $\chi'(T)$ below the crossover temperature (fig. S8).

a pronounced 1D character (*21*), which may be relevant to the understanding of the unusual NFL behavior of $YbNi_4P_2$. Because adjacent Yb chains are shifted by $c/2$, Yb^{3+} ions experience a crystalline electrical field (CEF) that is rotated by 90° from chain to chain, although the symmetry on each Yb site is the same. This causes frustration between neighboring chains, which together with strong electron-electron correlations may enhance the strength of quantum fluctuations.

We first focus on stoichiometric $YbNi_4P_2$. The strongest evidence of the FM order is provided by the real-part of the ac-susceptibility, $\chi'(T)$, measured both with field $H \parallel c$ ($\chi'_\parallel$) and $H \perp c$ ($\chi'_\perp$) (Fig. 1A). Although the CEF anisotropy results in a much larger ground-state moment along the c direction than in the (a,b) plane, the FM ordering is perpendicular to c. As T is lowered toward $T_C = (0.150 \pm 0.005)$ K, $\chi'_\perp$ increases sharply up to a value of ~200×10^{-6} m^3/mol (≈4 in SI units). Below T_C, a pronounced increase of the imaginary part $\chi''_\perp$ of $\chi_{ac}(T)$ ($H \perp c$) is observed, representative of energy dissipation (fig. S7). In the same T region, $\chi'_\perp$ is only weakly T-dependent, possibly because the coercive field of the magnetic hysteresis is very small (*21*). $\chi'_\parallel$ increases with decreasing T, following a $T^{-0.66}$ power law below 10 K down to T_C, where it saturates at a value of ~17×10^{-6} m^3/mol. This anisotropy of the susceptibility suggests that the spins align perpendicularly to the c direction, and $\chi'_\parallel$ measures the transversal fluctuations. Clear kinks are seen at T_C in the resistivity curves $\rho_\perp(T)$ and $\rho_\parallel(T)$, with current j perpendicular (Fig. 1B) or parallel (Fig. 1C) to the c axis. Below T_C, the resistivity follows a T^n power law, with $n_\perp = 2.9$ and $n_\parallel = 2.5$, very likely because of charge-carrier scattering from FM magnons. As in the polycrystals, the resistivity is quasi-T-linear just above T_C. Such a NFL property, when observed over a sufficiently extended temperature range, is commonly considered a hallmark of quantum criticality (*1*, *2*, *15*, *16*). Because of the relatively low Kondo temperature, $T_K \approx 8$ K, a proper $\rho \sim T^n$ behavior is observed only within the rather narrow temperature window $0.15\ \text{K} \le T \le 1$ K. In this range, $\rho = \rho_0 + AT^n$, with $n_\perp = 1$ and $n_\parallel = 1.2$, and the coefficient $A_\perp = 21.6$ microhm·cm/K is

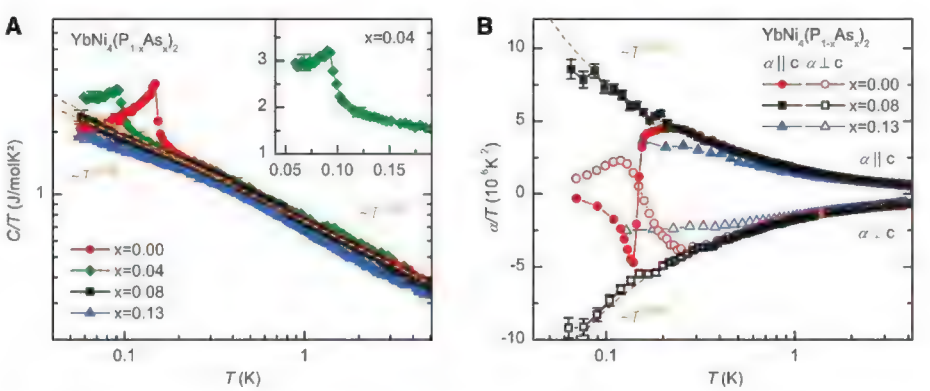


Fig. 3. Tuning T_C to zero by As substitution. (**A**) Specific heat $C(T)$ divided by T of $YbNi_4(P_{1-x}As_x)_2$ for indicated values of x at $H = 0$. With increasing x, the FM transition temperature is tuned to zero. For $x = 0.08$, C/T follows a $T^{-0.43}$ power law with a deviation below 0.25 K to $T^{-0.36}$ (orange dashed lines). (**B**) Linear thermal-expansion coefficient $\alpha(T)$ divided by T of $YbNi_4(P_{1-x}As_x)_2$ with $x = 0$, 0.08, and 0.13 measured in zero field along ($\alpha \parallel c$) and perpendicular ($\alpha \perp c$) to the c axis. Here, the orange dashed lines emphasize the $T^{-0.65}$ and $-T^{-0.66}$ power-law behavior of α/T versus T along the crystallographic c and a directions, respectively. For $YbNi_4(P_{0.87}As_{0.13})_2$, the coefficient $\alpha(T)/T$ tends to saturate below 1 K, indicating noncritical behavior.

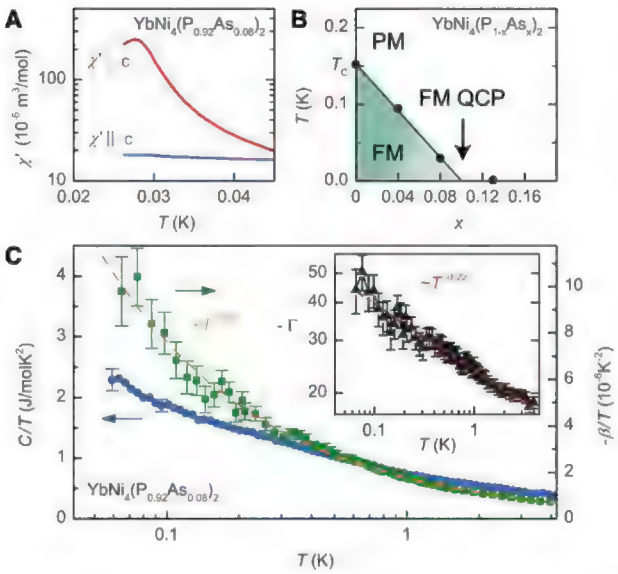


Fig. 4. Quantum critical behavior in $YbNi_4(P_{1-x}As_x)_2$. (**A**) Low-temperature ac-susceptibility for $x = 0.08$, measured at $H = 0$, where a rounded maximum is visible at $T_C = 0.028$ K in $\chi'_\perp(T)$. (**B**) $T - x$ phase diagram: The FM QCP is located at $x_c \approx 0.1$. (**C**) Temperature dependence of C/T and β/T for the $YbNi_4(P_{0.92}As_{0.08})_2$ sample, which is the sample closest to the FM QCP. Both quantities diverge with decreasing T as $C(T)/T \propto T^{-0.43}$ ($\propto T^{-0.36}$ below 0.25 K) and $\beta(T)/T \propto T^{-0.64}$. (Inset) Temperature dependence of the absolute value of the dimensionless Grüneisen ratio $\Gamma = V_{mol}/\kappa_T \cdot \beta/C$ for $x = 0.08$ [using the isothermal compressibility $\kappa_T = 5.3 \times 10^{-12}$ Pa^{-1} of $YbRh_2Si_2$ (*33*)] which reveals the presence of quantum critical fluctuations at the FM QCP. The dashed red line emphasizes the $T^{-0.22}$ divergence of $\Gamma(T)$ in zero field.

much larger than $A_\parallel = 3.8$ microhm·cm/K$^{1.2}$. This indicates a Fermi velocity that is much higher along the Yb chains than within the (a,b) plane, which is in accordance with the anisotropic Fermi surface (*21*).

Additional thermodynamic evidence of the phase transition at T_C and of the in-plane FM order is presented in Fig. 2, A and B, in which the T and H dependences of the specific heat are shown together with the result of the single-ion Kondo model (*25*). At $H = 0$, a sharp λ-type phase transition in C/T versus T is observed, associated with a small change of the molar 4*f*-derived entropy (1% of Rln2, with $R = 8.31$ J/molK). A magnetic field applied parallel to the c axis shifts the transition temperature to lower T (Fig. 2A), whereas for $H \perp c$, the crossover to the polarized state is shifted to higher temperatures (Fig. 2B). This is corroborated by the H dependence of $\chi'(T)$ (fig. S8, A and B): For $H \perp c$, a sharp peak appears already at very low fields, where the system directly enters the polarized state, and then broadens at higher fields. The resulting magnetic phase diagrams are displayed in Fig. 2, C and D. These observations point to in-plane FM order in $YbNi_4P_2$, although neutron-diffraction experiments are required to determine the exact magnetic structure, which very likely is not collinear.

At $H = 0$, $C(T)/T$ increases below 6 K, following a $T^{-0.43\pm0.02}$ power law over a notably wide range down to 0.2 K, where classical thermal fluctuations associated with the phase transition set in. A similar power law [$C(T)/T \propto T^{-0.4}$] has been observed at the field-induced QCP of $YbRh_2Si_2$, an AFM system with strong FM fluctuations (*15*). For $YbNi_4P_2$, a pronounced λ-peak at T_C indicates a second-order phase transition. This, and the distinct NFL properties above T_C, suggest that this compound is located in the close vicinity of a QCP, where $T_C \to 0$ smoothly.

To approach the QCP, we used arsenic substitutions of $x = 0.04$, 0.08, and 0.13 in $YbNi_4(P_{1-x}As_x)_2$. The QCP is achieved at $x_c \approx 0.1$, as demonstrated by calorimetric and dilatometric results in Fig. 3, A and B, and in the phase diagram of Fig. 4B. For $x = 0.04$, the phase transition remains FM and second order (Fig. 3A, green diamonds); in fact, $C(T)/T$ of $YbNi_4(P_{0.96}As_{0.04})_2$ has been measured at $H \perp c$, and the transition was found to be shifted to higher T with increasing H (fig. S6). For $x = 0.08$, $C(T)/T$ increases continuously down to 0.06 K, and no phase transition is detected. $C(T)/T$ can be described by the same T-dependence as for $YbNi_4P_2$ down to 0.25 K, where a slight deviation from this power law (to $T^{-0.36}$) is visible (Fig. 3A, orange dashed lines). Pronounced changes of the T-dependence of $C(T)/T$ and of the Grüneisen ratio $\Gamma(T)$ were observed in $YbRh_2Si_2$ below 0.3 K (*5*, *15*, *20*); in our case, this deviation results in only a minor change of $\Gamma(T)$ (figs. S12 and S13). Measurements of $\chi'(T)$ down to 0.02 K picked up a transition at 0.028 K, which is similar to the behavior observed in $YbNi_4P_2$ (Fig. 4A and fig. S9). This implies that, also for $x = 0.08$, the transition is FM and second order. For $x = 0.13$, no phase transition is observed, and $C(T)/T$ versus T seems to follow the same T dependence down to 0.1 K, but with smaller absolute values. This indicates that the same power-law divergence observed in $C(T)/T$ for the pure system is retained across x_c.

The observed NFL features suggest the presence of strong quantum fluctuations (*12*, *13*, *26*–*29*) but do not directly imply the existence of a FM QCP, that these fluctuations become critical at $x_c \approx 0.1$ (*13*, *29*). One key experiment to provide direct evidence of quantum critical fluctuations is the combined measurement of the volume-thermal-expansion coefficient $\beta = \alpha_\parallel + 2\alpha_\perp$ (with $\alpha_\parallel$ and $\alpha_\perp$ being the linear expansion coefficients along c and a, respectively) and the specific heat of the same crystal, with x as close as possible to x_c (*19*, *20*). The Grüneisen ratio $\Gamma(T)$ at x_c should diverge as $T^{-1/\nu z}$ with parameters ν, the spatial correlation-length exponent, and z, the dynamical exponent (*19*).

Whereas for $YbNi_4P_2$, a change of sign in $\alpha(T)/T$ is clearly visible at T_C (Fig. 3B), as expected at a second-order phase transition (*30*), for $YbNi_4(P_{0.92}As_{0.08})_2$, no sign change can be seen down to 0.06 K, in accordance with $T_C = 0.028$ K. Upon cooling, the ratios $\alpha_\parallel/T$ and $\alpha_\perp/T$ diverge with similar exponents, yielding $\beta(T)/T \propto T^{-0.64\pm0.02}$. Although this is consistent with 3D FM spin fluctuations (*19*), the power-law divergence of the Sommerfeld coefficient $C(T)/T$ is at odds with the expected logarithmic behavior, $C(T)/T \propto \log(1/T)$, for an itinerant 3D FM QCP (*19*). For $x = 0.13$, $\alpha(T)/T$ saturates below 1 K, indicating the crossover to a heavy Fermi liquid phase (*31*).

The key evidence for the FM QCP in $YbNi_4(P_{1-x}As_x)_2$ is summarized in Fig. 4C. At $x = 0.08$—very close to x_c—not only do both $C(T)/T$ and $\beta(T)/T$ diverge with power-law exponents $T^{-0.43}$ ($T^{-0.36}$ below 0.25 K) and $T^{-0.64}$, respectively, but the dimensionless Grüneisen ratio diverges as $\Gamma(T) \propto T^{-0.22\pm0.04}$ (Fig. 4C, inset). This strongly indicates the existence of a QCP at $x_c \approx 0.1$. The critical exponent $\lambda = -0.22$ rules out an itinerant QCP scenario because it leads to $\nu z \approx 5$, which is much larger than for itinerant FM systems ($\nu = 1/2$, $z = 3$).

The unexpected power-law exponents in all thermodynamic quantities, indicating the presence of strong FM quantum critical fluctuations, require new theoretical concepts beyond the itinerant spin-fluctuation theory (*2*). These concepts must take into account the localized nature of the Yb-4*f* states in the presence of very strong spin-orbit coupling when compared with the CEF. Further, the Kondo coupling of the 4*f* states to the conduction electrons must be considered. Isothermal measurements of the Hall coefficient on $YbNi_4(P_{1-x}As_x)_2$ with $x > \sim 0.1$ under hydrostatic pressure would greatly help to investigate the exciting possibility that the FM QCP concurs with a localization-delocalization transition of the Yb-4*f* states, similar to what was reported for $YbRh_2Si_2$ at its AFM QCP (*5*, *16*, *17*). The extent to which the FM instability in As-doped $YbNi_4P_2$ is affected by the 1D character of the electronic structure must also be clarified.

References and Notes

1. P. Coleman, A. J. Schofield, *Nature* **433**, 226 (2005).
2. H. Löhneysen, A. Rosch, M. Vojta, P. Wölfle, *Rev. Mod. Phys.* **79**, 1015 (2007).
3. G. R. Stewart, *Rev. Mod. Phys.* **73**, 797 (2001).
4. A. Yeh *et al.*, *Nature* **419**, 459 (2002).
5. J. Custers *et al.*, *Nature* **424**, 524 (2003).
6. D. Belitz, T. R. Kirkpatrick, T. Vojta, *Phys. Rev. Lett.* **82**, 4707 (1999).
7. A. V. Chubukov, C. Pépin, J. Rech, *Phys. Rev. Lett.* **92**, 147003 (2004).
8. G. J. Conduit, A. G. Green, B. D. Simons, *Phys. Rev. Lett.* **103**, 207201 (2009).

9. D. Belitz, T. Kirkpatrick, T. Vojta, *Rev. Mod. Phys.* **77**, 579 (2005).
10. E. Hassinger, D. Aoki, G. Knebel, J. Flouquet, *J. Phys. Soc. Jpn.* **77**, 073703 (2008).
11. S. S. Saxena *et al.*, *Nature* **406**, 587 (2000).
12. M. Brando *et al.*, *Phys. Rev. Lett.* **101**, 026401 (2008).
13. T. Westerkamp *et al.*, *Phys. Rev. Lett.* **102**, 206404 (2009).
14. D. Belitz, T. R. Kirkpatrick, http://arxiv.org/abs/1204.0873.
15. P. Gegenwart, Q. Si, F. Steglich, *Nat. Phys.* **4**, 186 (2008).
16. Q. Si, F. Steglich, *Science* **329**, 1161 (2010).
17. H. Pfau *et al.*, *Nature* **484**, 493 (2012).
18. S. J. Yamamoto, Q. Si, *Proc. Natl. Acad. Sci. U.S.A.* **107**, 15704 (2010).
19. L. Zhu, M. Garst, A. Rosch, Q. Si, *Phys. Rev. Lett.* **91**, 066404 (2003).
20. R. Küchler *et al.*, *Phys. Rev. Lett.* **91**, 066405 (2003).
21. C. Krellner *et al.*, *New J. Phys.* **13**, 103014 (2011).
22. C. Krellner, C. Geibel, *J. Phys. Conf. Ser.* **391**, 012032 (2012).
23. J. Spehling *et al.*, *Phys. Rev. B* **85**, 140406 (2012).
24. A. L. Cornelius, J. S. Schilling, D. Mandrus, J. D. Thompson, *Phys. Rev. B* **52**, R15699 (1995).
25. K. D. Schotte, U. Schotte, *Phys. Lett. A* **55**, 38 (1975).
26. N. T. Huy *et al.*, *Phys. Rev. B* **75**, 212405 (2007).
27. R. P. Smith *et al.*, *Nature* **455**, 1220 (2008).
28. S. Jia *et al.*, *Nat. Phys.* **7**, 207 (2011).
29. C. Pfleiderer, P. Böni, T. Keller, U. K. Rössler, A. Rosch, *Science* **316**, 1871 (2007).
30. M. Garst, A. Rosch, *Phys. Rev. B* **72**, 205129 (2005).
31. P. Thalmeier, B. Lüthi, in *Handbook on the Physics and Chemistry of Rare Earths*, K. A. Gschneidner Jr., L. Eyring, Eds. (North-Holland, New York, 1991), vol. 14, chap. 96, pp. 225–341.
32. Materials and methods are available as supplementary materials on *Science* Online.
33. J. Plessel *et al.*, *Phys. Rev. B* **67**, 180403 (2003).

Acknowledgments: We are indebted to D. Belitz, A. V. Chubukov, P. Coleman, R. Daou, S. Friedemann, M. Garst, P. Gegenwart, F. M. Grosche, S. Kirchner, G. G. Lonzarich, B. Schmidt, Q. Si, P. Thalmeier, M. Vojta, T. Vojta and S. Wirth for useful discussions and M. C. Schuman for language editing. Part of the work was supported by the DFG Research Unit 960 "Quantum Phase Transitions."

Supplementary Materials

www.sciencemag.org/cgi/content/full/339/6122/933/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S13
References (*34–50*)

21 September 2012; accepted 17 December 2012
10.1126/science.1230583

Living Crystals of Light-Activated Colloidal Surfers

Jeremie Palacci,[1]* Stefano Sacanna,[1] Asher Preska Steinberg,[2] David J. Pine,[1] Paul M. Chaikin[1]

Spontaneous formation of colonies of bacteria or flocks of birds are examples of self-organization in active living matter. Here, we demonstrate a form of self-organization from nonequilibrium driving forces in a suspension of synthetic photoactivated colloidal particles. They lead to two-dimensional "living crystals," which form, break, explode, and re-form elsewhere. The dynamic assembly results from a competition between self-propulsion of particles and an attractive interaction induced respectively by osmotic and phoretic effects and activated by light. We measured a transition from normal to giant-number fluctuations. Our experiments are quantitatively described by simple numerical simulations. We show that the existence of the living crystals is intrinsically related to the out-of-equilibrium collisions of the self-propelled particles.

Self-organization often develops in thermal equilibrium as a consequence of entropy and potential interactions. However, there are a growing number of phenomena where order arises in driven, dissipative systems, far from equilibrium. Examples include "random organization" of sheared colloidal suspensions (*1*) and rods (*2*), nematic order from giant-number fluctuations in vibrated rods (*3*), and phase separation from self-induced diffusion gradients (*4*). Biological (*5–7*) and artificial active particles (*8–11*) also exhibit swarm patterns that result from their interactions (*12–15*).

In order to study active, driven, collective phenomena, we created a system of self-propelled particles where the propulsion can be turned on and off with a blue light. This switch provides rapid control of particle propulsion and a convenient means to distinguish nonequilibrium activity from thermal Brownian motion. Further, the particles are slightly magnetic and can be stabilized and steered by application of a modest magnetic field. Our system consists of an

[1]Department of Physics, New York University, 4 Washington Place, New York, NY 10003, USA. [2]Department of Physics and Chemistry, Brandeis University, Waltham, MA 02453, USA.

*To whom correspondence should be addressed. E-mail: jp153@nyu.edu

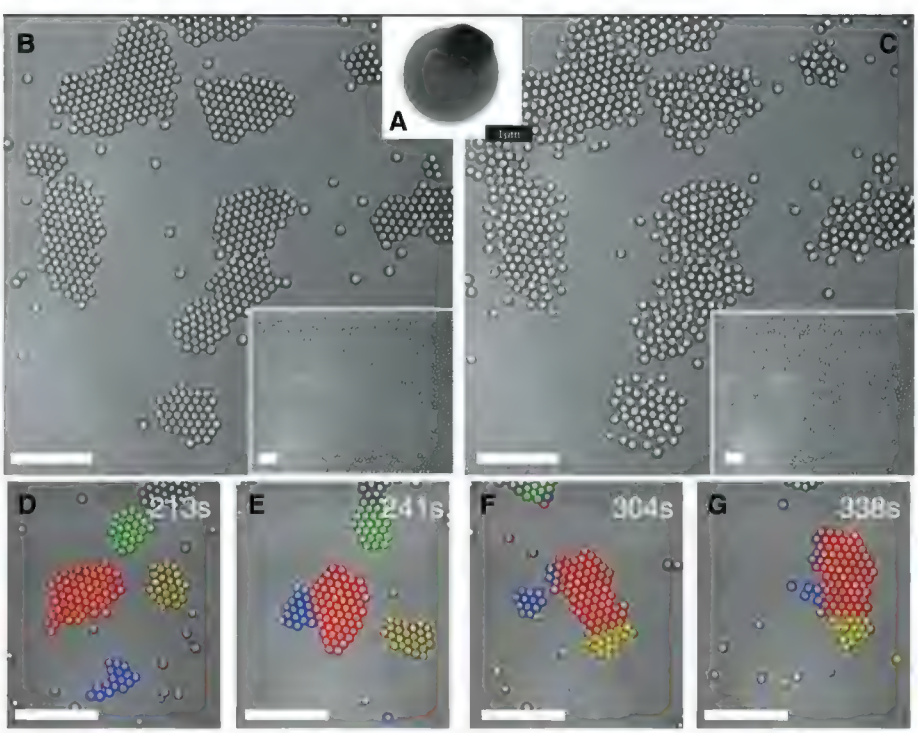


Fig. 1. (**A**) Scanning electron microscopy (SEM) of the bimaterial colloid: a TPM polymer colloidal sphere with protruding hematite cube (dark). (**B**) Living crystals assembled from a homogeneous distribution (inset) under illumination by blue light. (**C**) Living crystals melt by thermal diffusion when light is extinguished: Image shows system 10 s after blue light is turned off (inset, after 100 s). (**D** to **G**) The false colors show the time evolution of particles belonging to different clusters. The clusters are not static but rearrange, exchange particles, merge (D→F), break apart (E→F), or become unstable and explode (blue cluster, F→G). For (B) to (G), the scale bars indicate 10 μm. The solid area fraction is $\Phi_s \approx 0.14$.

active bimaterial colloid. A polymer sphere, 3-methacryloxypropyl trimethoxysilane (TPM), encapsulates most of a canted antiferromagnetic hematite cube (*16*), but with part exposed to the solvent (Fig. 1A). The particles are immersed in a basic solution (pH ~ 8.5) containing hydrogen peroxide [0.1 to 3% weight/weight (w/w)], 5 mM tetramethylammonium hydroxide, and 3.4 mM sodium dodecyl sulfate (SDS) (*16*). Under normal bright-field illumination, the colloids are at equilibrium with the solvent and thus sediment toward the bottom glass surface of the observation cell. When illuminated through the microscope objective with blue-violet light ($\lambda \sim 430$ to 490 nm), the particles exhibit self-propulsion (movie S1). The motion, with a velocity up to 15 μm/s, only takes place at the cell surface, whether it is the bottom, the vertical side walls, or the cell top. Individual particles undergo a random walk with a persistence length determined by the reorientation time $\tau_r = 8.0 \pm 1.5$ s, consistent with Stokes-Einstein rotational diffusion.

In equilibrium, with no blue light, the particles diffused and were disordered (Fig. 1B, inset). At surface area fractions $\Phi_s \gtrsim 7\%$, cooperative behavior of the light-activated colloids began to emerge. Crystallites started to form in the sample 25 s after the light was turned on. An image of the crystals after 350 s is shown in the main part of Fig. 1B. Immediately after the light was extinguished, the crystallites began "dissolving" because of thermal diffusion (Fig. 1C); after 100 s, there was no trace of the crystals (Fig. 1C, inset). Although the particles formed organized crystallite structures under illumination, those structures were not static (movie S2). The crystallites actively translated and rotated, collided, joined, and split (Fig. 1, D to G). These "living crystals" reached a dynamic steady state of creation and self-destruction. They did not grow to incorporate all available particles, as would be the case for an equilibrium system with attractive interactions. We measured an average cluster size of ~35 particles, which did not seem to depend on $\Phi_s > 10\%$. The lifetime of the crystals was finite and broadly distributed. The typical time for a cluster to change its size by 50% is 100 ± 75 s. Fluctuations in the local number N of particles follows a power law $\Delta N \sim N^{\alpha}$. There is a transition at $\Phi_S \sim 7\%$ from normal $\alpha = 1/2$ to giant fluctuations $\alpha \cong 0.9$, in line with recent predictions for disordered clusters in a system of polar isotropic active particles (*17*) and observed in a granular system (*18*).

To understand the mechanisms involved in the self-propulsion and crystallization, we performed a series of experiments on the separate components of our composite colloid, the hematite cube, and the polymer sphere. First, we attached a hematite cube to a glass substrate and immersed it in our solution of surfactants, buffer, and H_2O_2. Then, 1.5-μm-diameter colloidal tracer particles of polystyrene, silica, or TPM were added. The tracers were observed to diffuse randomly under normal bright-field illumination. When illuminated with blue light, however, the tracers all moved toward the immobilized hematite cube (movie S3), converging on it from all directions, as indicated by the cartoon in Fig. 2A. This observation rules out advection, because advective fluid flow must have zero divergence. Therefore, the motion of the colloids toward the hematite particle must be caused by a phoretic motion (*19*) induced by some gradient generated by the cube. Under blue-light illumination, hematite catalyzes the exothermic chemical decomposition of H_2O_2, creating thermal and chemical (H_2O_2 and O_2) gradients. Heating studies of the system suggest that diffusiophoresis is more important than thermophoresis in our system. The motion of the tracers toward the cube can be quantified by

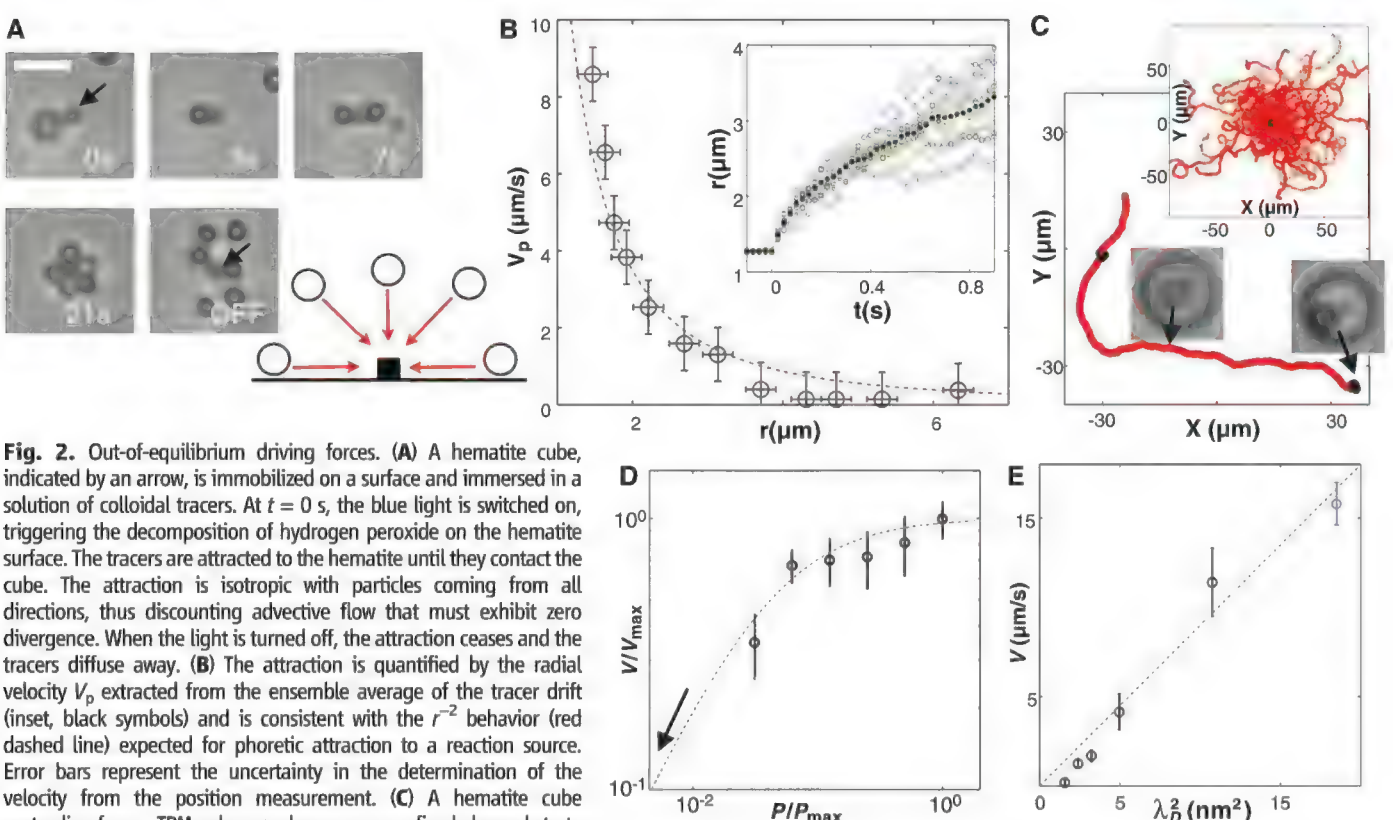


Fig. 2. Out-of-equilibrium driving forces. (**A**) A hematite cube, indicated by an arrow, is immobilized on a surface and immersed in a solution of colloidal tracers. At $t = 0$ s, the blue light is switched on, triggering the decomposition of hydrogen peroxide on the hematite surface. The tracers are attracted to the hematite until they contact the cube. The attraction is isotropic with particles coming from all directions, thus discounting advective flow that must exhibit zero divergence. When the light is turned off, the attraction ceases and the tracers diffuse away. (**B**) The attraction is quantified by the radial velocity V_p extracted from the ensemble average of the tracer drift (inset, black symbols) and is consistent with the r^{-2} behavior (red dashed line) expected for phoretic attraction to a reaction source. Error bars represent the uncertainty in the determination of the velocity from the position measurement. (**C**) A hematite cube protruding from a TPM polymer sphere moves on fixed glass substrate when exposed to blue light (red part of trace) and diffuses when the light is off (black part of trace). Initially, with no light, the hematite cube is oriented randomly (image, right) but rotates and faces downward toward the glass substrate when the light is turned on (image, left). The particle then surfs on the osmotic flow it induces between the substrate and itself. (Inset) A superposition of the trajectories of many particles with their origins aligned. (**D**) The particle velocity V increases with light intensity P and follows Michaelis-Menten law (red dashed line). The black arrow indicates the point of zero velocity for $P = 0$. (**E**) The particle velocity is also strongly dependent on the Debye length λ_D of the system and asymptotically follows the $V \propto \lambda_D^2$ scaling expected for osmotic mechanisms (red dashed line). The Debye length is varied by adding NaCl to the buffer solution except for the blue symbol, for which the SDS surfactant is suppressed to reach higher λ_D. The error bars in (D) and (E) are given by the standard deviation of the velocity measured for 10 to 20 different particles.

monitoring their position versus time and calculating their velocity as a function of distance from the particle (Fig. 2B). The dashed line through the data is a fit to A/r^2 (where A is a fitting parameter and r the center-to-center distance between the tracer and the hematite particle), consistent with a diffusive concentration profile $C \propto C_\infty(1 - B/r)$.

Phoresis and osmosis are complementary interfacial phenomena: In a gradient, a free colloid will exhibit a phoretic migration, whereas a fixed surface of the same material will exhibit an osmotic flow at its surface in the opposite direction (*19*). Therefore, a particle phoresing to the right has an osmotic flow at its surface to the left. Just as a silica colloid is attracted to a hematite cube, a free hematite particle is attracted to a stationary silica surface. Therefore, when we add free cubes to a sample cell, the silica surface of the cover slip attracts the cubes. Indeed, we observed that hematite cubes were quickly drawn to the glass substrate as soon as blue light was turned on.

Unexpectedly, once on the glass substrate, the hematite cubes continue to move on the glass surface when illuminated with blue light. Naïvely, one might expect the cubes to remain stationary, because the osmotic flow on the cover slip surface is away from the cube and ideally should be symmetric. However, the symmetry is broken either by imperfections on the cube or spontaneously by an instability where the motion of the cube induces different gradients fore and aft. Thus, in a solution of free hematite cubes we see attraction of the cubes to the surface followed by self-propulsion of the cubes surfing on the substrate when the light is turned on.

When using a suspension of our composite particles, a hematite cube in a TPM sphere, we observed a similar scenario. When illuminated with blue light, the composite particle reoriented so that the exposed hematite sat on the glass substrate (Fig. 2C, insets) and then began to move at speeds comparable to that of the hematite alone. Figure 2C shows the trajectory of a single composite particle, with the light turned on then off, whereas the inset shows a superposition of many trajectories with their origins aligned. The self-propelled motions are isotropic and diffusive with a persistence length (15 to 100 μm) determined by the rotational diffusion time and the velocity of the particle. The velocity of the particles depends weakly on the H_2O_2 concentration but strongly on the light intensity and the Debye screening length. The velocity versus light intensity, P, follows Michaelis-Menten law (*20*) behavior characteristic of a catalytic reaction (Fig. 2D). Figure 2E suggests that the composite particle velocity asymptotes to a quadratic behavior with Debye length λ_D (*21*), a behavior expected from osmotic effects within a Debye length of a surface where the driving force $\propto \lambda_D$ and the drag force is $\propto$ velocity/λ_D (*19*, *22*).

If we now consider a solution of composite particles activated by light, two effects have to be taken into account: (i) the collisions between our self-propelled particles surfing on the osmotic flow they set up and (ii) the phoretic attraction between the particles. In order to see whether these effects explain the formation of our living crystals, we performed simulations guided by our experimentally determined parameters.

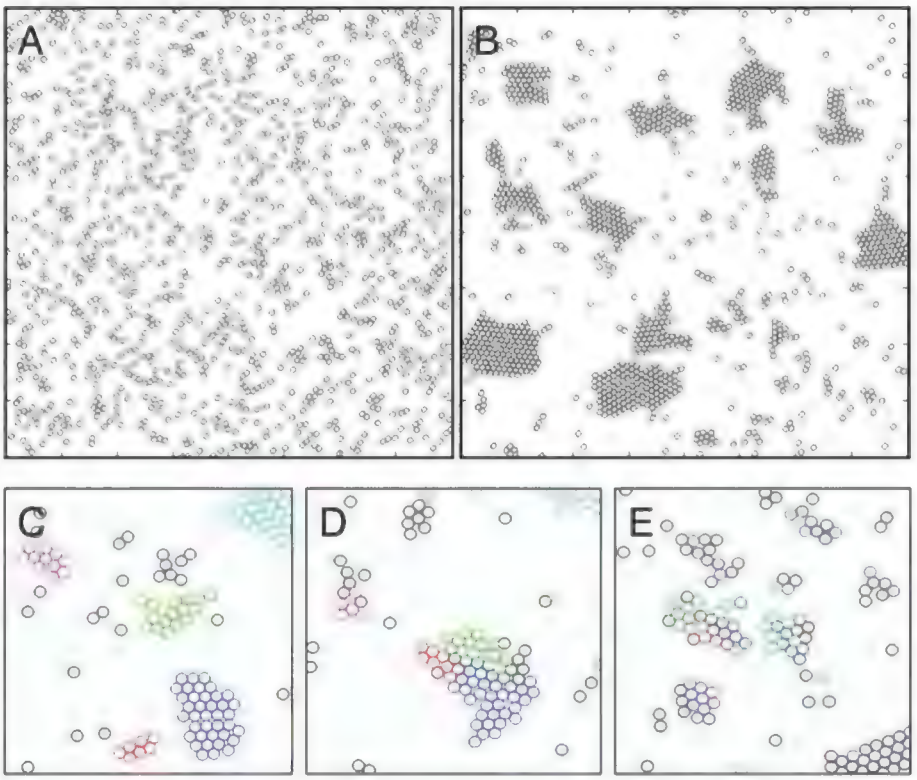


Fig. 3. Numerical simulations of self-propelled disks coupled by a phoretic attraction. Simulation parameters were defined to mimic the experimental conditions. (**A**) Starting from a homogeneous distribution (**B**) ($\sim t = 8\sim\tau_r$), the disks assemble in mobile crystalline clusters with faceted edges. (**C** to **E**) The false colors show the time evolution of particles belonging to different clusters. The crystals are mobile [(C) to (E)], can merge [(C) and (D)], and can break apart or dissolve [(D) and (E)]. This minimal model reproduces the experimental dynamics of the living crystals and demonstrates the relevance of the proposed mechanism. We used parameters consistent with the experimental conditions of Fig. 1: Φ_s = 0.14, $\tilde{\tau}_r$ = 16 and $\tilde{A}$ = 0.87 [see (*16*) for a definition of the reduced parameters, indicated by tildes].

We considered a minimal numerical model (*16*) in which the self-propelled colloids are represented by self-propelled hard disks that move with a constant velocity in a direction that changes randomly on a time scale τ_r determined by rotational diffusion. We modeled the phoretic attraction between particles as a pairwise attraction between nearby particles consistent with the phoretic velocity (red line in Fig. 2B). If a displacement makes two disks overlap, the particles are separated by moving each one half the overlap distance along their center-to-center axis. We tried various approximations to account for the effect of (hydrodynamic) lubrication forces in the crystals, for example, increase of the apparent viscosity, and found little qualitative difference.

In Fig. 3 and movie S4, we present the results of simulations in which the attractive phoretic effects are taken into account. For our experimental conditions $\Phi_s \sim 3$ to 20% and ~ 300 to 1500, the simulations reproduce the crystallite formation as well as the size and lifetime of the crystallites quite well. The simulations (Fig. 4, A and B) also capture the transition observed experimentally from normal to giant fluctuations of number, above a critical density $\Phi_S^C \simeq 7\% \pm 1\%$. If we turned off the pairwise phoretic attraction, we observed large spatiotemporal fluctuations of density with normal number fluctuations, $\alpha \sim 0.5$ (Fig. 4B). The formation of clusters that grow and decay is recovered at much higher particle concentration, $\Phi_s > \sim 35$ to 45% (fig. S2).

Our understanding of the living crystals comes from the idea that active particles undergo diffusive motion with a large persistence length when not in "contact" but slow down, translate, and diffuse more slowly when they are in contact. The slowing down results from the inability of particles to penetrate their neighbors when encountered.

Several recent papers have pointed out that such density-dependent diffusion can lead to giant fluctuations, clustering, and phase separation in nonequilibrium systems (*17*, *23*–*26*), however, at much larger surface density. The formation of noncrystalline clusters of active particles has been reported with bacteria coupled through short-range depletion interaction (*27*) and Janus particles with chemical attraction (*28*).

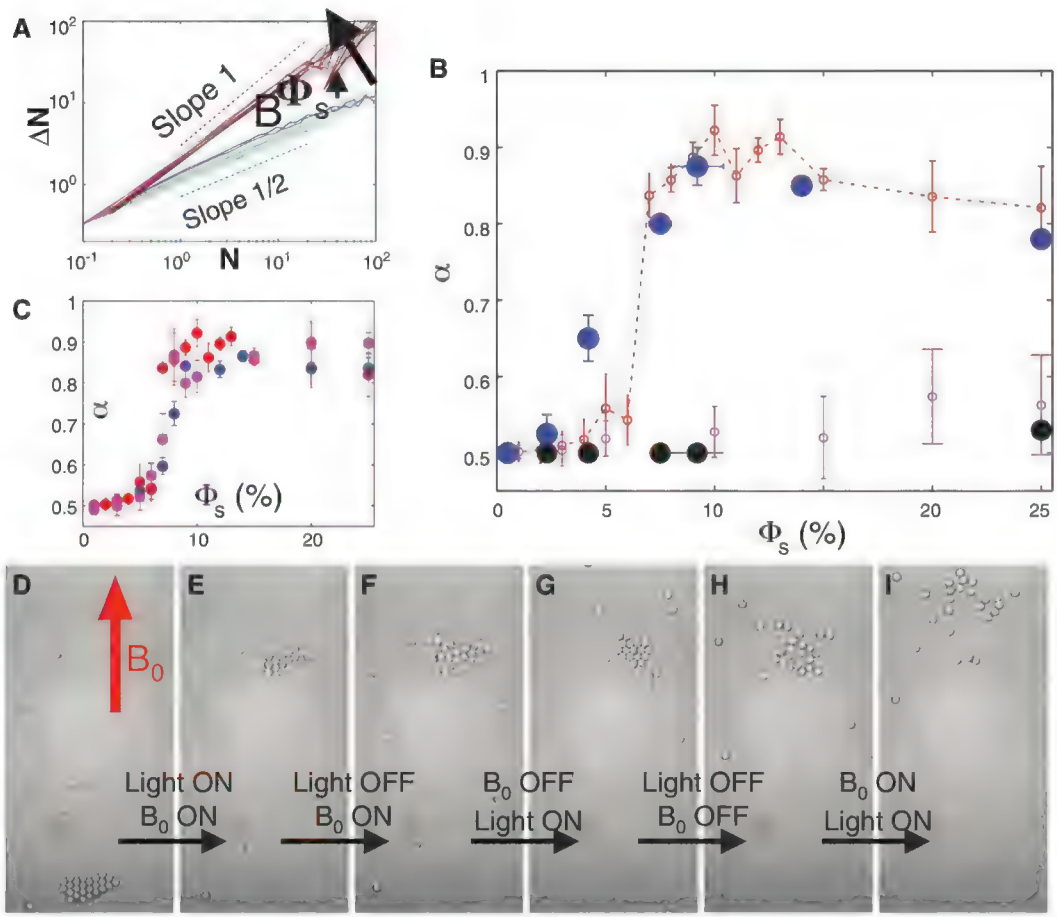


Fig. 4. (**A**) Number fluctuations measured in the simulations for varying Φ_s values, in the range 1 to 15% for $N = 600$ particles [$\tilde{\tau}_r = 16$ and $\tilde{A} = 0.87$ from the experiment, see (*16*)]. The system exhibits a transition from normal to giant fluctuations for $\Phi_s^c \sim 7\%$. (**B**) Scaling α of the number fluctuations, $\Delta N \propto N^{\alpha}$, for various Φ_s values measured in the experiment at equilibrium (black symbols), under activation by the light (blue symbols), and in the simulations with ($\tilde{A} = 0.87$, red open symbols) and without ($\tilde{A} = 0$, magenta open symbols) attraction. We observed normal fluctuations, $\alpha = 1/2$, at equilibrium. The driven system exhibited a transition from normal fluctuations, $\alpha = 1/2$, to giant-number fluctuations, $\alpha \sim 0.9$, at $\Phi_s \sim 7\%$ in both the experiment and the simulations. The slight decay of the exponent observed in experiments and simulations is a finite size effect. Error bars indicate the standard deviation in the measurement of the exponent. (**C**) Scaling α of the number fluctuations for $N = 1000$ (blue symbols), $N = 600$ (red), and $N = 400$ (magenta) particles in the simulations. For $N = 400$, the curve after $50\tau_0$ (square) collapses with $35\tau_p$ (circles), showing that the scaling is steady. (**D** to **I**) Investigating the crystal mechanism. We used an external magnetic field $B_0 \sim 1$ mT to orient the particles and direct their motion. The red arrow is the orientation of B_0. (D and E) The magnetic field is turned on, and the light is on; the crystal is self-propelled in the direction of the magnetic field, and crystal breakup is suppressed. (F) The light is turned off, and the magnetic field B_0 is left; the crystal dissolves. (G) The magnetic field is turned off, and the light is turned on; particles collide, and the crystal re-forms. (H) The light is turned off, and the magnetic field remains off; the crystal dissolves. (I) The magnetic field is turned on first, and then the light is turned on. The particles all move in the the field direction; they do not collide and do not crystallize.

We always observed the intermittent formation and breakage of the large crystals. This differs from the equilibrium nucleation and growth of a crystal of attractive colloids or the asymptotic formation of a single cluster from an assembly of self-propelled disks reported by Fily and Marchetti (*17*) (at higher number and density of disks than in our experiment).

In order to investigate the underlying mechanisms, we took advantage of the magnetic properties of the embedded hematite cube. Under an external magnetic field ($B \sim 1$ mT), the alignment of the hematite slighty tilts the orientation of the particle, and the self-propulsion proceeds in a direction parallel to the field, suppressing the rotational diffusion.

We test two crucial aspects of our scenario in Fig. 4, D to I, and movie S5. We suggest that the crystal breakup depends on the velocity redirection by rotational diffusion of the particles in the crystal. If the direction of all the particles are aligned by using an external magnetic field, then the crystal will not break up. This is shown in Fig. 4, D and E, and movie S5. Also note the suppression of the breakup in the magnetically steered crystal in movie S6. We also claim that collisions are required for crystal formation. In Fig. 4F we turned off the light and the crystal dissolved. With the magnetic field off, we turned on the light, the particles collided, and the crystal re-formed (Fig. 4G). We turned off the light and redissolved the crystal in Fig. 4H. We then first turned on the magnetic field and then the light (Fig. 4I). The particles all moved in the same direction and did not collide, and there was no crystallization.

We have demonstrated a form of self-assembly from nonequilibrium driving forces leading to living crystals with complex dynamics. The osmotically driven motion and steric hindrances lead to the formation of dynamic aggregates. The introduction of a small attractive interaction, in this case arising from phoresis, orders the aggregates into periodic crystals at low surface fraction. Phoretic and osmotic effects in our system can conveniently be switched on and off by light. Rotational diffusion of the particles reorients their motions, leading to a finite persistence length, crystal evaporation and breakup, and finite size and lifetime of the crystals. The use of active particles and nonequilibrium forces for directed self-assembly opens a new area for design and production of novel and moving structures. The fact that they can be turned on and off with visible light adds control to this system, as does the ability to use external magnetic steering.

References and Notes

1. L. Corté, P. M. Chaikin, J. P. Gollub, D. J. Pine, *Nat. Phys.* **4**, 420 (2008).
2. A. Franceschini, E. Filippidi, E. Guazzelli, D. J. Pine, *Phys. Rev. Lett.* **107**, 250603 (2011).
3. V. Narayan, S. Ramaswamy, N. Menon, *Science* **317**, 105 (2007).
4. M. E. Cates, D. Marenduzzo, I. Pagonabarraga, J. Tailleur, *Proc. Natl. Acad. Sci. U.S.A.* **107**, 11715 (2010).
5. J.-F. Joanny, S. Ramaswamy, *Nature* **467**, 33 (2010).
6. A. Cavagna *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **107**, 11865 (2010).
7. M. Polin, I. Tuval, K. Drescher, J. P. Gollub, R. E. Goldstein, *Science* **325**, 487 (2009).
8. R. Dreyfus *et al.*, *Nature* **437**, 862 (2005).
9. W. F. Paxton *et al.*, *J. Am. Chem. Soc.* **126**, 13424 (2004).
10. J. R. Howse *et al.*, *Phys. Rev. Lett.* **99**, 048102 (2007).
11. J. Palacci, C. Cottin-Bizonne, C. Ybert, L. Bocquet, *Phys. Rev. Lett.* **105**, 088304 (2010).
12. D. Saintillan, M. J. Shelley, *Phys. Fluids* **20**, 123304 (2008).

13. H. P. Zhang, A. Be'er, E.-L. Florin, H. L. Swinney, *Proc. Natl. Acad. Sci. U.S.A.* **107**, 13626 (2010).
14. T. Vicsek, A. Czirók, E. Ben-Jacob, I. Cohen, O. Shochet, *Phys. Rev. Lett.* **75**, 1226 (1995).
15. G. Grégoire, H. Chaté, *Phys. Rev. Lett.* **92**, 025702 (2004).
16. Materials and methods are available as supplementary materials on *Science* Online.
17. Y. Fily, M. C. Marchetti, *Phys. Rev. Lett.* **108**, 235702 (2012).
18. J. Deseigne, O. Dauchot, H. Chaté, *Phys. Rev. Lett.* **105**, 098001 (2010).
19. J. Anderson, *Annu. Rev. Fluid Mech.* **21**, 61 (1989).
20. L. Michaelis, M. Menten, *Biochem. Z.* **49**, 333 (1913).
21. J. Israelachvili, *Intermolecular and Surface Forces* (Elsevier, Amsterdam, 2011).
22. R. Golestanian, T. B. Liverpool, A. Ajdari, *Phys. Rev. Lett.* **94**, 220801 (2005).
23. J. Bialké, T. Speck, H. Löwen, *Phys. Rev. Lett.* **108**, 168301 (2012).
24. S. R. McCandlish, A. Baskaran, M. F. Hagan, *Soft Matter* **8**, 2527 (2012).
25. F. D. C. Farrell, M. C. Marchetti, D. Marenduzzo, J. Tailleur, *Phys. Rev. Lett.* **108**, 248101 (2012).
26. J. Tailleur, M. E. Cates, *Phys. Rev. Lett.* **100**, 218103 (2008).
27. J. Schwarz-Linek *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **109**, 4052 (2012).
28. I. Theurkauff, C. Cottin-Bizonne, J. Palacci, C. Ybert, L. Bocquet, *Phys. Rev. Lett.* **108**, 268303 (2012).

Acknowledgments: We thank J. Layné, K. Hanson, A. Grosberg, R. Dreyfus, E. Lerner, A. Baskaran, and L. Bocquet for fruitful discussions. This work was supported by the Materials Research Science and Engineering Centers program of the NSF under award number DMR-0820341 and by the U.S. Army Research Office under grant award no. W911NF-10-1-0518. We acknowledge partial support from the NASA under grant award NNX08AK04G.

Supplementary Materials

www.sciencemag.org/cgi/content/full/339/6122/936/DC1
Materials and Methods
Figs. S1 and S2
References (*29*, *30*)
Movies S1 to S6

11 September 2012; accepted 12 December 2012
10.1126/science.1230020

Global Patterns of Groundwater Table Depth

Y. Fan,[1]* H. Li,[1] G. Miguez-Macho[2]

Shallow groundwater affects terrestrial ecosystems by sustaining river base-flow and root-zone soil water in the absence of rain, but little is known about the global patterns of water table depth and where it provides vital support for land ecosystems. We present global observations of water table depth compiled from government archives and literature, and fill in data gaps and infer patterns and processes using a groundwater model forced by modern climate, terrain, and sea level. Patterns in water table depth explain patterns in wetlands at the global scale and vegetation gradients at regional and local scales. Overall, shallow groundwater influences 22 to 32% of global land area, including ~15% as groundwater-fed surface water features and 7 to 17% with the water table or its capillary fringe within plant rooting depths.

The groundwater table is an undulating surface between the oxygenated soils and the water-saturated aquifers below. Where shallow, it can influence land surface in multiple ways. For instance, shallow groundwater provides base-flow to rivers and lakes maintaining aquatic ecosystems in dry periods (*1*, *2*). Shallow groundwater tables impede drainage and create water-logged soil conditions defining wetland habitats (*2*, *3*). A shallow water table supplies root-zone soil water and photosynthetic production in droughts through capillary rise (*4*–*7*). To understand the influence of groundwater on terrestrial ecosystems, a regional and global assessment of water table depth (WTD) is required. However, WTD monitoring and reporting are highly variable across the globe and often limited to developed regions, and despite the many large-scale syntheses of groundwater conditions (*8*–*13*), there remains the lack of a global water table map.

We present global observations of WTD at 1,603,781 well sites compiled from government archives and published literature (*14*) (Fig. 1). Where government data are absent, we searched the literature for published data, which is the main source for Africa and Asia, where large data gaps still remain. The mean is shown where time series are available, but >90% of the sites have only one reading (at different times). The observed WTD varied from land surface to >200 m, but the distribution peaks at 2 to 7 m (Fig. 2). The 2- to 7-m peak reflects sampling bias; observations are made for resource monitoring where humans settle (excluding large swamps and deserts) and where the water table is lowered by pumping or drainage (figs. S1 and S2).

At first glance, there is a lack of coherent patterns; shallow WTD is found in both humid and arid climates and at both low and high elevations. However, some global trends do emerge. First, it is common for WTD to be <5 m. Second, the shallowest WTD (<0.25 m) is more common in humid climates (e.g., eastern and northern North America), pointing to some level of climate control. Third, shallow WTD is also found in arid valleys (e.g., the southwest United States), suggesting terrain-driven moisture convergence that overrides climate divisions. However, the large spatial gaps and temporal inconsistency prohibit further insights and invite a mechanistic model to elucidate the interplays among the multiple drivers at various scales.

To fill in observation gaps and infer patterns and process controls, we used a previously developed groundwater model forced by present-day climate, terrain, and sea level (*3*) (fig. S3). The model simulates vertically integrated lateral groundwater movement among grid cells at 30 arc-second (~1 km) grid spacing to resolve multiscale convergence within computational limits (*14*). We assume that, first, the permeability of continental crust drops exponentially with depth and the rate depends on land slope (*14*), so that flat valleys accumulate thick sediments (fig. S4C). Second, at any location, there is only one water table (neglecting local, perched aquifers). Third, groundwater pumping, irrigation, and drainage are not represented by the global land models providing recharge here. Thus, the model neglects local complexity and human influence and only captures the broad-scale and natural patterns that are our primary focus.

The model water table (Fig. 3) seems deeper than observations in arid regions because the observations favor valleys and oases (shallow WTD) where humans live, as seen when plotting only grid cells with observations (fig. S6). We evaluated the model by examining the residual (model-observation difference) over different continental regions (figs. S7 to S11), which suggests that the minimally simple model offers a realistic, albeit broad-scaled, sketch of global water table condition that can shed light on spatial patterns and processes.

The WTD distribution (Fig. 2) differs substantially between observations and the model results. The model suggests a much higher fraction of WTD <1 m because it includes surface water features (rivers, lakes, and inundated wetlands) fed by groundwater discharge (WTD ≤ 0) that are not sampled by wells, and few observations are in the vast boreal and interior swamps. At the deeper end, the model suggests a higher fraction of WTD > 30 m because few observations are in large deserts. The higher occurrence of observations in the 2- to 30-m range reflects observational bias.

The model offers a globally continuous but simpler (climate equilibrium, neglecting local geology) view of WTD at its natural states (without pumping or drainage). At the global scale, sea level is the dominant driver; WTD is shallow along the margins of all continents in all climates, but the ribbon of shallow WTD is wider where flat coastal plains meet the sea. This ribbon correlates well with the coastal wetlands of the world, some extending far inland across flat

[1]Department of Earth and Planetary Sciences, Rutgers University, New Brunswick, NJ 08854, USA. [2]Nonlinear Physics Group, Faculty of Physics, Universidade de Santiago de Compostela, Galicia, Spain.

*To whom correspondence should be addressed. E-mail: yingfan@rci.rutgers.edu

coastal plains and up lowland valleys. Sea-level forcing on land hydrology is well known (although not implemented in large-scale models); coastal wetlands have migrated up and down the continental shelves following sea-level change through past glacial-interglacial shifts (*3*, *15*).

At the regional scale, climate emerges as the primary driver. In mid- and low latitudes, regions of deep WTD correspond to regions of low recharge; the great deserts of the world stand out. So do the tropical swamps with high recharge (Amazon and Orinoco). In the boreal region,

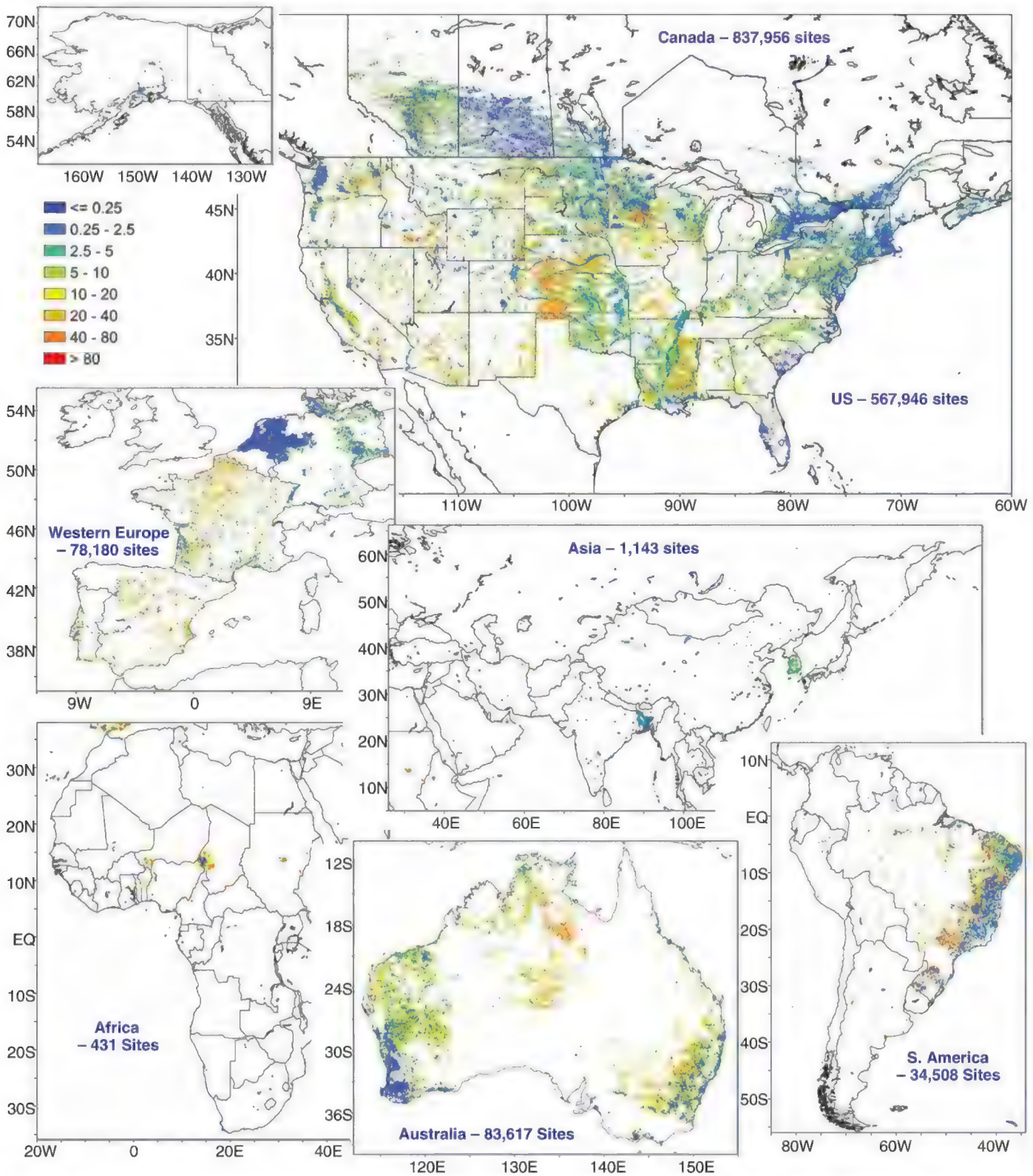


Fig. 1. Observed water table depth (m) compiled from government archives and literature (1,603,781 sites). Temporal mean is shown where time series are available.

climate control takes the form of frozen ground, where the thin thawed layer limits drainage. The large patches of shallow water table here coincide with Pan-Arctic peatlands of the world. The thin active layer is compounded by flat, low terrain, pointing to regional topographic gradient as another driver. Topographic control is also evident in warmer climates of central Amazonia and along the large swath of flat lowlands in South America from the Pantanal through the Chaco to the Pampas; large inland freshwater wetlands characterize the landscape here.

At more local scales, terrain signals dominate. Lateral convergence from high to low grounds creates a texture in WTD that overrides climate boundaries. Larger examples are arid basins where groundwater convergence from surrounding mountains maintains valley ecosystems (oases) otherwise absent (figs. S12B, S14B, and S15B). Smaller examples are river valleys etched into plateaus in semiarid climate or in humid climate with pronounced dry seasons, creating a marked gradient in water availability from valleys to ridge tops (fig. S13B). The low recharge is concentrated into a small fraction of the landscape by topography, forming discharge zones that harbor closed lakes, oases, gallery forests, and riparian wetlands, many designated as Ramsar sites of international ecological importance (table S2 and figs. S12 to S16). It is known that valleys are wetter spots and that topography is a powerful predictor of hydrologic states.

Where solar energy is not limiting, global distributions of vegetation are strongly aligned with moisture gradients indicated by annual or seasonal rainfall (*16*, *17*). However, rainfall alone cannot explain desert oases, the latter fed by groundwater due to topography-driven lateral flow that redistributes the moisture surplus or deficit across the landscape (*3*, *18*), giving patterns to WTD that defy rainfall (shallow in deserts and wide-ranging under same climate). We highlight the utility of WTD as an additional (to rainfall) moisture regulator and ecological filter with two examples, the geographic distribution of wetlands at the global scale and species differentiation and adaptation at the landscape scale.

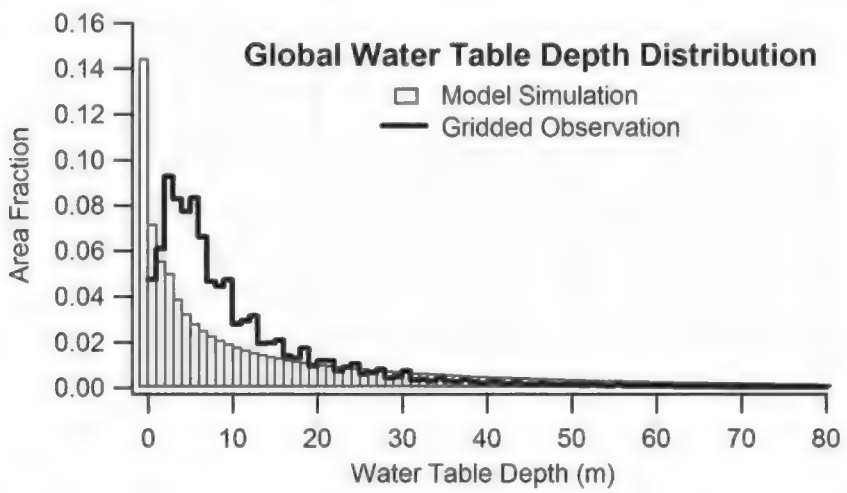


Fig. 2. Global distribution of simulated (gray) and observed (bold black) water table depth (observations first gridded into 30 arc second cells).

Wetlands are habitats of vegetation adapted to water-logged soils. Small seasonal wetlands can result from local rainfall, but extensive and prolonged water-logging must be caused by shallow water tables. This is the reason that the water table map resembles the wetland map (*3*) and that the areal extents of shallow water tables and wetlands are highly correlated (fig. S17). Within a given climate regime and history, species differentiation aligns with local environmental gradients (*17*). A well-articulated gradient is the topography from valley to ridge (*19*, *20*) spanning decameters to kilometers. Because valleys are enduring features of the landscape, and because hill-to-valley groundwater convergence is slow and steady, the shallow water table in the valley offers a stable moisture source from below, creating predictable moisture gradients, particularly in dry regions or seasons. A large body of evidence suggests that WTD is a powerful niche differentiator that sorts vegetation and microbial species into anoxia-tolerant at shallow to drought-tolerant at deep water table (*5*–*7*, *19*–*29*), and for a given species, WTD regulates biomass production and drives physiological adaptation such as rooting characteristics (*6*, *30*–*34*).

The model gives a first-order estimate of global land area likely affected by shallow groundwater (*14*); ~15% is covered by lakes (excluding the large lakes in Fig. 3), rivers, and inundated wetlands fed by persistent groundwater discharge (water table rising above land surface, WTD ≤ 0), ~2% by less frequently inundated wetlands ($0 < \text{WTD} \leq 0.25$ m), and 5 to 15% with WTD or its capillary fringe within the rooting depth of upland plants (*14*), adding to 22 to 32% of global land area. These results suggest a widespread and

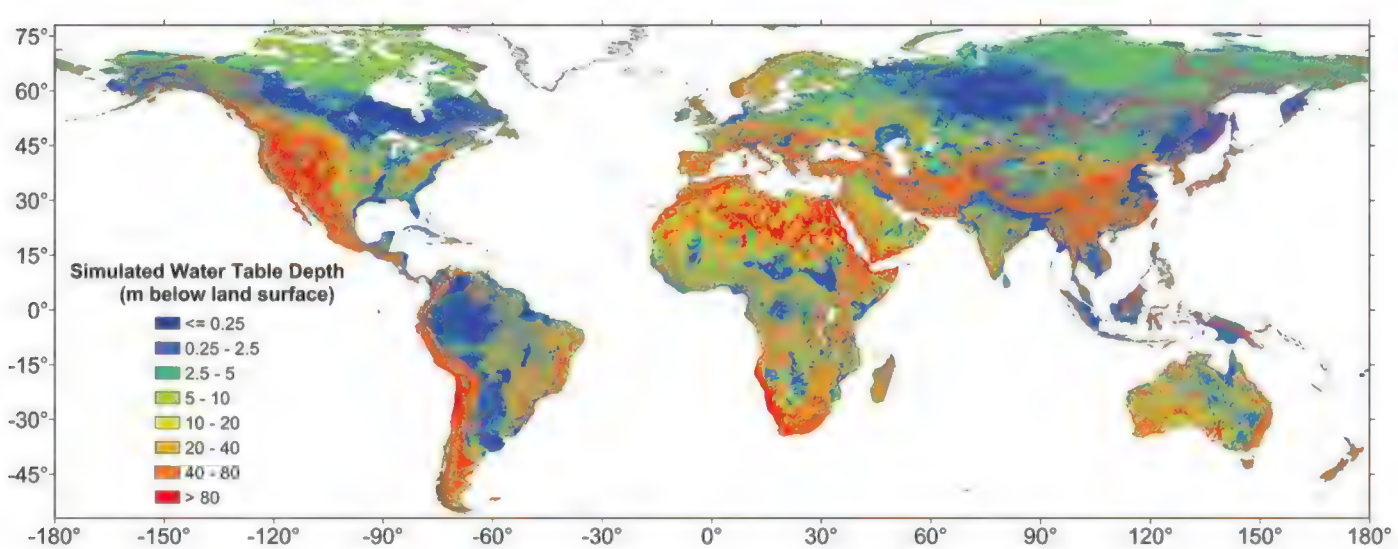


Fig. 3. Simulated water table depth (m) at 30 arc-sec grid (~1 km) constrained by observations in Fig. 1.

structured influence of groundwater on land hydrology and ecosystems and highlight the need for larger efforts to improve observing and modeling large-scale groundwater processes in the context of earth system dynamics.

References and Notes

1. T. C. Winter, J. W. Harvey, O. L. Franke, W. A. Alley, "Ground water and surface water: A single resource" (USGS Circular 1139, U.S. Government Printing Office, Denver, CO, 1998).
2. G. Miguez-Macho, Y. Fan, *J. Geophys. Res.* **117**, (D15), D15113 (2012a).
3. Y. Fan, G. Miguez-Macho, *Clim. Dyn.* **37**, 253 (2011).
4. G. Miguez-Macho, Y. Fan, *J. Geophys. Res.* **117**, (D15), D15114 (2012b).
5. D. C. Martre, D. F. Scott, C. Colvin, *Water S.A.* **25**, 137 (1999).
6. F. Orellana, P. Verma, S. P. Loheide II, E. Daly, *Rev. Geophys.* **50**, RG3003 (2012).
7. D. R. Rossatto, L. de Carvalho Ramos Silva, R. Villalobos-Vega, L. S. L. Sternberg, A. C. Franco, *Environ. Exp. Bot.* **77**, 259 (2012).
8. P. Döll, K. Fiedler, *Hydrol. Earth Syst. Sci.* **12**, 863 (2008).
9. M. Giordano, *Annu. Rev. Environ. Resour.* **34**, 153 (2009).
10. T. L. Gleeson *et al.*, *Geophys. Res. Lett.* **38**, L02401 (2011).
11. T. R. Green *et al.*, *J. Hydrol.* **405**, 532 (2011).
12. A. M. MacDonald, H. C. Bonsor, B. E. O. Docharaigh, R. G. Taylor, *Environ. Res. Lett.* **7**, 024009 (2012).
13. R. G. Taylor *et al.*, *Nat. Clim. Change* **3**, nclimate1744 (2012).
14. Materials and methods are available as supplementary material on *Science* Online.
15. H. F. Faure, R. C. Walter, D. R. Grant, *Global Planet. Change* **33**, 47 (2002).
16. B. A. Hawkins *et al.*, *Ecology* **84**, 3105 (2003).
17. H. Kreft, W. Jetz, *Proc. Natl. Acad. Sci. U.S.A.* **104**, 5925 (2007).
18. E. G. Jobbágy, M. D. Nosetto, P. E. Villagra, R. B. Jackson, *Ecol. Appl.* **21**, 678 (2011).
19. B. M. J. Engelbrecht *et al.*, *Nature* **447**, 80 (2007).
20. C. E. T. Paine, K. E. Harms, J. Ramos, *J. Trop. Ecol.* **25**, 171 (2009).
21. K. A. Dwire, J. B. Kauffman, J. E. Baham, *Wetlands* **26**, 131 (2006).
22. A. J. Elmore, J. F. Mustard, S. J. Manning, *Ecol. Appl.* **13**, 443 (2003).
23. J. Grogan, J. Galvao, *Acta Amazon.* **36**, 483 (2006).
24. F. M. R. Hughes, *J. Biogeogr.* **15**, 127 (1988).
25. S. Jirka *et al.*, *J. Veg. Sci.* **18**, 183 (2007).
26. R. Pélissier, S. Dray, D. Sabatier, *Plant Ecol.* **162**, 143 (2002).
27. M. A. Sobrado, *J. Trop. Ecol.* **26**, 215 (2010).
28. J. C. Stromberg, R. Tiller, B. Richter, *Ecol. Appl.* **6**, 113 (1996).
29. A. A. Bobrov, D. J. Charman, B. G. Warner, *Protist* **150**, 125 (1999).
30. S. K. Arndt, A. Kahmen, C. Arampatsis, M. Popp, M. Adams, *Oecologia* **141**, 385 (2004).
31. M. R. Bakker, L. Augusto, D. L. Achat, *Plant Soil* **286**, 37 (2006).
32. J. L. Carter, D. A. White, *Tree Physiol.* **29**, 1407 (2009).
33. H. S. Mishra, T. R. Rathore, V. S. Tomar, *Irrig. Sci.* **18**, 117 (1999).
34. J. Stave, G. Oba, A. B. Eriksen, I. Nordal, N. C. Stenseth, *For. Ecol. Manage.* **212**, 367 (2005).

Acknowledgments: We thank many individuals for making and providing observations (full acknowledgement is in the supplementary materials) and P. Döll for providing recharge estimates. Funding comes from U.S. National Science Foundation (NSF-AGS-1045110 and NSF-OCE-10409088), U.S. Environmental Protection Agency (EPA-STAR-RD834190), European Commission FP7 (GLOWASIS), and a Rutgers University Board of Trustees grant (651201). Computation used the Extreme Science and Engineering Discovery Environment (XSEDE) supported by U.S. National Science Foundation (NSF-OCI-1053575) and the Climate Simulation Laboratory at NCAR's Computational and Information Systems Laboratory, sponsored by the National Science Foundation and other agencies. Observations, code, and model results are accessible at https://glowasis.deltares.nl/thredds/catalog/opendap/opendap/Equilibrium_Water_Table/catalog.html.

Supplementary Materials

www.sciencemag.org/cgi/content/full/339/6122/940/DC1
Supplementary Text
Figs. S1 to S17
Tables S1 to S3
References (*35*–*228*)
Databases S1 to S3

10.1126/science.1229881

Direct Observations of Atmospheric Aerosol Nucleation

Markku Kulmala,[1]* Jenni Kontkanen,[1] Heikki Junninen,[1] Katrianne Lehtipalo,[1] Hanna E. Manninen,[1] Tuomo Nieminen,[1,14] Tuukka Petäjä,[1] Mikko Sipilä,[1] Siegfried Schobesberger,[1] Pekka Rantala,[1] Alessandro Franchin,[1] Tuija Jokinen,[1] Emma Järvinen,[1] Mikko Äijälä,[1] Juha Kangasluoma,[1] Jani Hakala,[1] Pasi P. Aalto,[1] Pauli Paasonen,[1] Jyri Mikkilä,[2] Joonas Vanhanen,[2] Juho Aalto,[3] Hannele Hakola,[4] Ulla Makkonen,[4] Taina Ruuskanen,[1] Roy L. Mauldin III,[1,5] Jonathan Duplissy,[1] Hanna Vehkamäki,[1] Jaana Bäck,[6] Aki Kortelainen,[7] Ilona Riipinen,[8] Theo Kurtén,[1,9] Murray V. Johnston,[10] James N. Smith,[7,11] Mikael Ehn,[1,12] Thomas F. Mentel,[12] Kari E. J. Lehtinen,[4,7] Ari Laaksonen,[4,7] Veli-Matti Kerminen,[1] Douglas R. Worsnop[1,4,7,13]

Atmospheric nucleation is the dominant source of aerosol particles in the global atmosphere and an important player in aerosol climatic effects. The key steps of this process occur in the sub–2-nanometer (nm) size range, in which direct size-segregated observations have not been possible until very recently. Here, we present detailed observations of atmospheric nanoparticles and clusters down to 1-nm mobility diameter. We identified three separate size regimes below 2-nm diameter that build up a physically, chemically, and dynamically consistent framework on atmospheric nucleation—more specifically, aerosol formation via neutral pathways. Our findings emphasize the important role of organic compounds in atmospheric aerosol formation, subsequent aerosol growth, radiative forcing and associated feedbacks between biogenic emissions, clouds, and climate.

Atmospheric aerosol formation [that is, the formation of molecular clusters and their growth to larger sizes (*1*, *2*)] has an important effect on aerosol particle number concentrations (*3*, *4*) and on climate through indirect radiative effects (*5*, *6*). To understand the initial steps of atmospheric aerosol formation, one must have detailed knowledge of the concentrations of neutral and charged clusters, their chemical composition, and gaseous compounds participating in their formation and growth. However, size-segregated measurements of sub–2-nm clusters are extremely rare, and until now, no one has taken comprehensive and simultaneous field measurements of charged and neutral clusters and their precursors (supplementary materials, section 3).

Recent technical developments make it possible to measure the concentrations and size distributions of ions, molecular clusters, and nanoparticles in the 1- to 2-nm mobility diameter range and to simultaneously obtain information about the chemical composition of these entities and their interactions with trace gases. Here, we present a comprehensive analysis of such measurements, conducted between 14 March and 16 May 2011, at the SMEAR II station (*7*) in Hyytiälä, southern Finland. We measured the total nanoparticle and ion concentrations, along with the concentrations of gaseous compounds, including sulfuric acid, volatile organic compounds, ammonia, amines, ozone, sulfur dioxide, and nitrogen oxides. The instruments we used to take our measurements are described in greater detail in the supplementary materials (sections 1.3.1 to 1.3.7).

We categorized each day of the measurement campaign as a "nucleation event day," a "non-event day," or an "undefined day" (table S7) (*8*). We determined the concentrations of nanoparticles and ions separately for six size classes between 0.9 and 2.1 nm (supplementary materials, section 1.2). For each size class, we calculated the concentration of nanoparticles originating from neutral formation pathways, N_n, from the relation $N_{tot} = N_{ions} + N_{rec} + N_n$, where N_{tot} is the total measured nanoparticle concentration in that size class, N_{ions} is the corresponding ion concentration, and N_{rec} is the estimated concentration of neutral particles originating from the recombination of

[1]Department of Physics, University of Helsinki, Finland. [2]Airmodus Oy, Helsinki, Finland. [3]SMEAR Station II, Hyytiälä, Finland. [4]Finnish Meteorological Institute, Finland. [5]University of Colorado at Boulder, Boulder, CO, USA. [6]Department of Forest Sciences, University of Helsinki, Finland. [7]University of Eastern Finland, Kuopio, Finland. [8]University of Stockholm, Stockholm, Sweden. [9]Department of Chemistry, University of Helsinki, Finland. [10]University of Delaware, Newark, DE, USA. [11]National Center for Atmospheric Research, Boulder, CO, USA. [12]Forschungszentrum Juelich, IEK-8, 52425 Juelich, Germany. [13]Aerodyne Research, Billerica, MA, USA. [14]Helsinki Institute of Physics, Helsinki, Finland.

*To whom correspondence should be addressed. E-mail: markku.kulmala@helsinki.fi

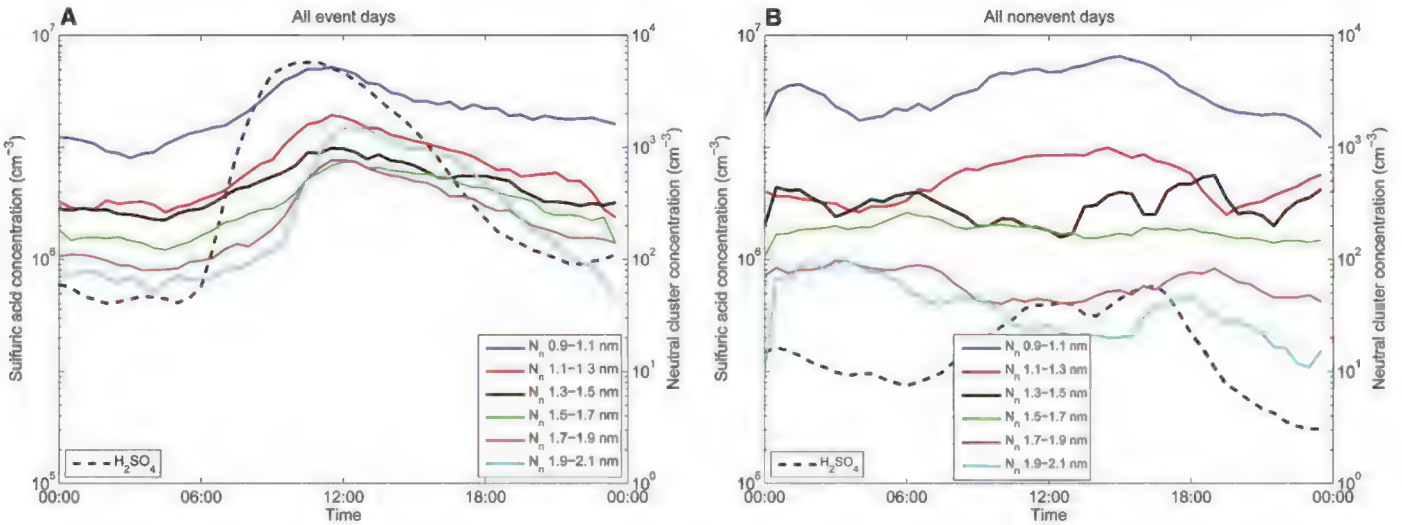


Fig. 1. Diurnal variations of the sulfuric acid concentration (left *y* axis) and the concentrations of neutral clusters in different size classes (right *y* axis) during median event days (**A**) and median nonevent days (**B**).

oppositely charged ions ending up in that size class (supplementary materials, section 1.2).

Excluding a few intermittent periods during the night, N_n exceeded N_{ions} and N_{rec} in all of the sub–2-nm size classes (fig. S9). The dominance of neutral nanoparticles over ions was particularly clear during the periods of active aerosol formation, characterized by the measurable formation of 3-nm particles and their further growth to larger sizes (see fig. S8A). Ions showed a clear concentration maximum below 1.5 nm and a steep concentration decrease above this size under all conditions (fig. S9A). Such a maximum corresponds to the size of stable ion clusters in the ion-induced nucleation theory. Overall, our observations indicate that the dynamics of sub–2-nm clusters were dominated by neutral pathways, with little influence due to the presence of ions. This finding is in line with the latest analysis of earlier field measurements conducted in various continental boundary-layer environments (*9*). Hereafter, we will focus our attention on the properties of neutral nanoparticles and call them "neutral clusters." In reality, the smallest nanoparticles include both clusters and large molecules or molecular complexes.

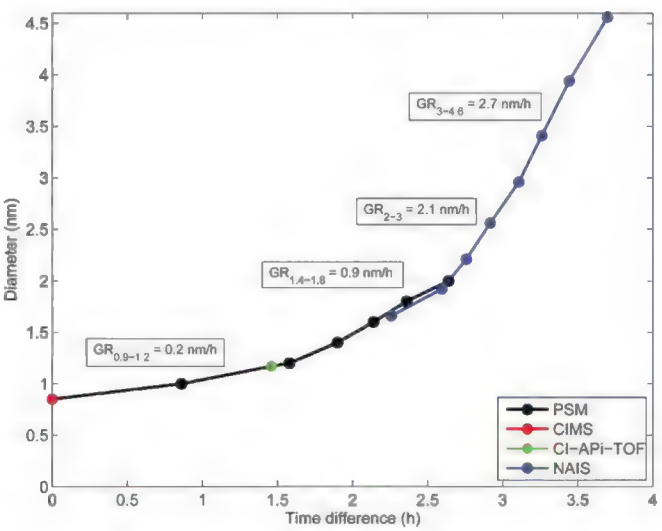


Fig. 2. Cluster size as a function of time during median nucleation event days. Here, we show data from normalized concentrations of sulfuric acid [from a chemical ionization mass spectrometer (CIMS)], size-segregated neutral clusters [from a particle size magnifier (PSM)], a cluster with a mass of 339.06 Th (from CI-APi-TOF), and ions/aerosol particles [neutral cluster and air ion spectrometer (NAIS)]. Times when the concentrations reached half of their maximum are plotted. Because sulfuric acid concentration rises first, the zero time is taken from the value of sulfuric acid. The used mobility diameters of a sulfuric acid molecule and a cluster with a mass of 339.06 Th are 0.85 and 1.17 nm, respectively. Growth rates (GR) for particles below 5-nm diameter were determined using the plotted data. The data with error bars are given in the supplementary materials (section 4 and fig. S15).

Figure 1 shows the median neutral cluster concentrations in different size classes and sulfuric acid concentrations during the nucleation event days (33 days) and nonevent days (19 days). The first thing we observe is the near-constant presence of neutral clusters throughout the sub–2-nm size range (see also fig. S8). This feature indicates continuous formation of neutral clusters and their subsequent growth up to a mobility diameter of at least 2 nm. Concentrations of neutral clusters were typically 10 times higher than concentrations of aerosol particles in the nucleation mode (3- to 12-nm diameter). The second important feature, discussed in more detail below, is the very different behavior of the clusters in three size regimes: (i) <1.2 ± 0.1 nm, (ii) from 1.2 ± 0.1 nm to 1.7 ± 0.2 nm, and (iii) >1.7 ± 0.2 nm. In the first size regime (the size classes 0.9 to 1.1 nm and 1.1 to 1.3 nm in Fig. 1), cluster concentrations and their diurnal behavior were very similar between the nucleation event and nonevent days. In the two other size regimes, cluster concentrations showed a clear daytime maximum during the nucleation event days, which was absent during the nonevent days. The cluster maxima peaked somewhat later than the sulfuric acid concentration. Cluster concentrations decreased rapidly with the increasing cluster size in the first two size regimes. In the third size regime, cluster concentration appeared to increase with the increasing cluster size around noon on nucleation event days (see also fig. S9A). The maximum concentrations of 1.9- to 2.1-nm neutral clusters were about a factor of 50 higher during the nucleation event days compared with the nonevent days.

Several groups have suggested that gaseous sulfuric acid is the main driver of daytime atmospheric aerosol formation (*1*, *10*–*12*). During our measurement campaign, we observed a clear association between the sulfuric acid concentration and the formation rate of 1.5-, 2.0-, and 3.0-nm clusters and aerosol particles (supplementary materials, section 5, and figs. S16 to S18). To look closer at this association, we determined the average growth rates of neutral sub–3-nm clusters during the periods of active aerosol formation (Fig. 2; see also supplementary materials, section 4). These rates were equal to 0.2, 0.9, and 2 nm h^{-1} in the mobility diameter ranges

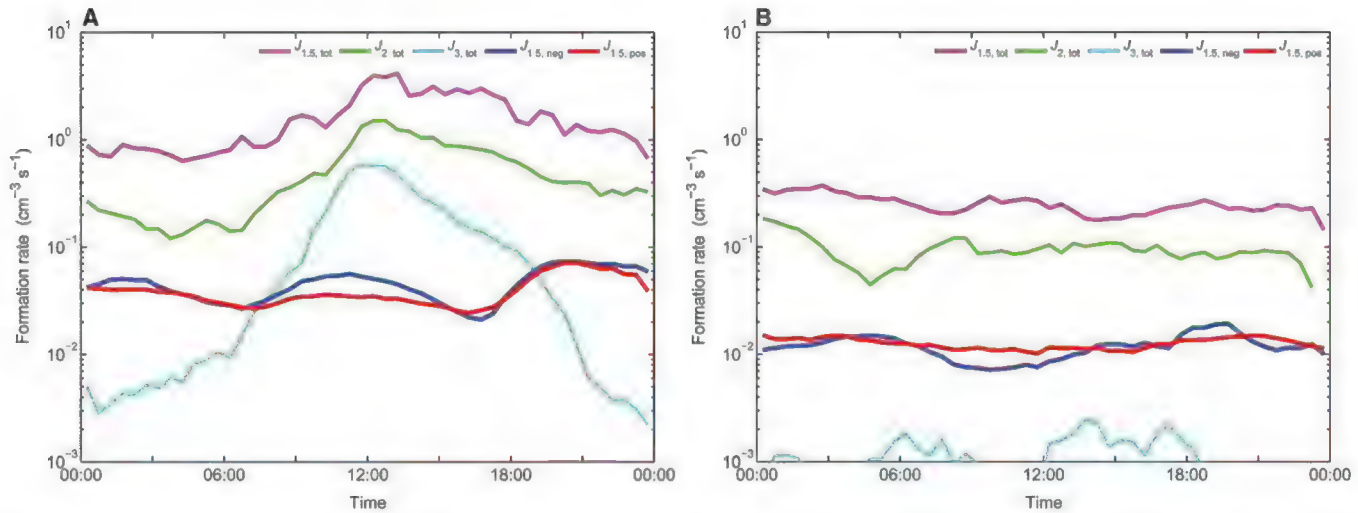


Fig. 3. Average diurnal cycle of the formation rates of 1.5-, 2.0-, and 3.0-nm atmospheric aerosol particles (or clusters) and 1.5-nm negative- and positive-ion clusters during (**A**) nucleation event days and (**B**) nonevent days. The particle-formation rates as a function of sulfuric acid concentration are given in the supplementary materials (section 5 and figs. S16 to S18).

of <1.2, 1.4 to 1.8, and 2 to 3 nm, respectively. An increase in cluster growth rates with increasing particle size in the sub–3-nm size range has recently been reported at two other measurement sites as well (*13*). Next, we calculated the maximum cluster growth rates due to sulfuric acid condensation over the same size ranges by allowing no sulfuric acid evaporation, and we found these rates to be 0.3 to 0.4, 0.6 to 0.9, and 0.5 to 0.6 nm h^{-1}, on average. Clearly, there was more than enough sulfuric acid vapor to explain the growth of the smallest clusters. Together with the very steep decline of cluster concentrations at these sizes, this finding suggests that neutral clusters in the first size regime (<1.2 ± 0.1 nm) both take up and evaporate vapor molecules constantly, the net effect being a slow cluster growth during the periods of active aerosol formation.

In the second size regime (1.2 ± 0.1 to 1.7 ± 0.2 nm), the concentration of sulfuric acid was, within experimental uncertainties, high enough to explain the observed cluster growth rates. On the other hand, quantum chemical investigations (*14*–*17*) predict that clusters containing only sulfuric acid (and water) cannot grow; rather, these clusters must be stabilized with amines, ammonia, or organic vapors (of these, amines are the strongest stabilizers) (*16*). Our measurements obtained by chemical ionization atmospheric pressure interface time-of-flight (CI-APi-TOF) mass spectrometer showed abundant presence of amines in the gas phase, and our APi-TOF measurements showed the existence of sulfuric acid–amine clusters (figs. S5 and S21 to S23). Small sulfuric acid–amine clusters have been identified in earlier laboratory experiments (*18*–*20*) and tentatively also in the atmosphere (*18*).

Clusters in the third size regime (>1.7 ± 0.2 nm) grew two to four times faster than what can be explained by sulfuric acid condensation. Furthermore, the strongest aerosol-formation events were characterized by a local minimum in the size space within this size regime (Fig. 1A and fig. S9A). Such a minimum is in accordance with Nano-Köhler theory (*21*), which describes the activation of neutral clusters by condensable vapors in a manner analogous to cloud condensation nuclei activation by water vapor during cloud formation. After activation, nanoparticles are expected to grow faster due to a decreasing Kelvin effect and, thus, an enhanced condensation flux. The most likely candidates for the additional compounds participating in the growth of sub–3-nm clusters and nanoparticles are organic vapors (*22*–*24*). The cluster activation in the size range of 1.5 to 2.0 nm could be caused by condensation of low-volatility organic vapors once they have overcome the Kelvin barrier (*21*, *25*) or by heterogeneous reactions between the clusters and organic vapors (*24*).

The CI-APi-TOF instrument detects a wide range of molecules and clusters in the 300- to 700-thomson (Th) range, mainly believed to result from highly oxidized organic molecules, similar to those recently found in the ambient APi-TOF measurements by Ehn *et al.* (*26*). The temporal behavior of the ions in this size range is highly variable, with some ions observed mainly at night and others during the day. During periods with high solar radiation, the most abundant cluster above 300 Th is typically detected at 339.06 ± 0.02 Th and has tentatively been identified as $C_{10}H_{15}N_2O_{11}^-$, with one or both N atoms resulting from clustering with nitrate and/or nitric acid inside the CI-APi-TOF instrument. Several other clusters in the 300- to 700-Th range show similar time behavior, and the cluster at 339 Th will be used to represent this group. Although the CI-APi-TOF detection and quantification of organics still needs detailed characterization, our results show that the diurnal variation in the signal of the 339-Th tracks 1.5- to 2.0-nm cluster concentrations even better than sulfuric acid (supplementary materials, section 6, and fig. S20). The total estimated concentration of highly oxidated organic molecules in the 300- to 450-Th range correlated positively with cluster concentrations in the second size regime during active aerosols formation (but not at any other time; see supplementary materials, table S5), indicating that these organic molecules will participate in atmospheric nucleation.

The net flux of growing clusters past a certain size, usually some fixed mobility diameter d, is defined as the formation rate J_d (*8*). Figure 3 shows the formation rates of ions and neutral nanoparticles over the diurnal cycle of an average particle-formation (nucleation) event day and nonevent day. During the nonevent days, the formation rates of 1.5-nm ions averaged to ~0.01 cm^{-3} s^{-1} and showed little diurnal variability. The formation rates of 1.5- and 2.0-nm neutral clusters exceeded 0.1 cm^{-3} s^{-1}, whereas the formation rates of 3-nm particles were mostly too low to be detected. During the days with active aerosol formation, the average formation rates of 1.5-nm ions approached 0.08 cm^{-3} s^{-1} around noon and late evening/early night. The average formation rates of neutral 1.5-nm particles exceeded those of 1.5-nm ions by about one order of magnitude and by up to a factor of 50 to 100 during the afternoon. The formation rates of 3-nm particles were measurable over the whole diurnal cycle of a typical nucleation event day, with values of J_3 approaching J_2 during a few hours around noon. This observation shows the importance of the enhanced growth rate from sub–2-nm size to 3 nm. The formation rate of 3-nm particles during event days was 100 to 1000 times higher than during nonevent days, with the corresponding difference at 1.5 or 2 nm being less than one order of magnitude.

The formation rate of 1.5-nm clusters is often called the atmospheric nucleation rate (*27*). However, from a thermodynamic point of view, nucleation requires overcoming the free energy

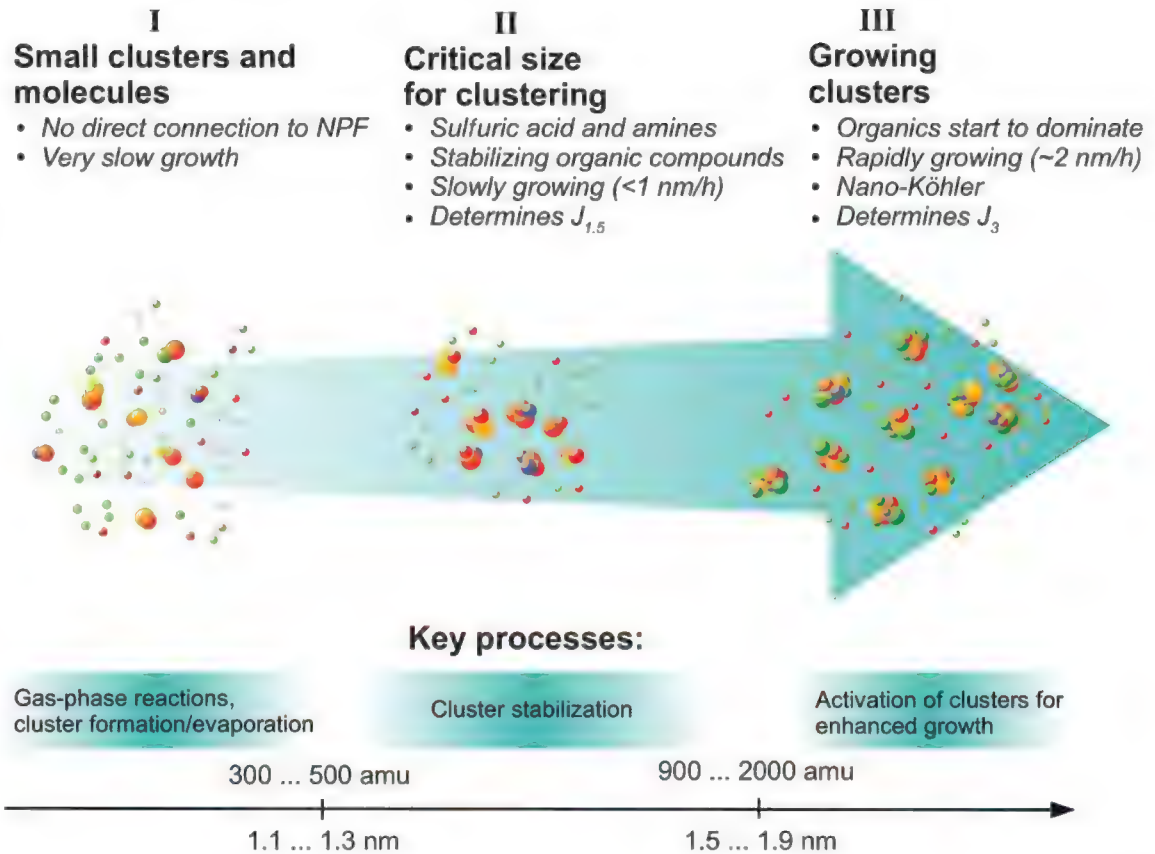


Fig. 4. Schematic description of main size regimes of atmospheric neutral clusters and the main processes related to those size ranges. In regime I (mobility diameter < 1.1 to 1.3 nm), the net cluster growth is very small, as clusters and molecules are continually formed and lost as a result of chemical reactions, vapor uptake, and evaporation. This regime has no direct connection to atmospheric nucleation. NPF, new particle formation. In regime II (1.1 to 1.3 nm to 1.5 to 1.9 nm) clusters grow in size by sulfuric acid condensation and are simultaneously being stabilized by amines, ammonia, or organic vapors. This regime is critical for the clustering process (atmospheric nucleation) and the formation rate of 1.5-nm neutral clusters. Additionally, gas-phase chemical reactions are important in this size range, and activation of clusters will begin. In regime III (>1.5 to 1.9 nm), clusters grow faster than can be explained by sulfuric acid concentration, especially during the periods of active aerosol formation when large numbers of 3-nm aerosol particles are being formed. This fast rate of growth is due to enhanced vapor uptake (most likely of oxidized organic vapors). The key process is activation, by which vapors can contribute to the enhanced growth. In addition, clustering plays a role, and condensable vapors are produced via chemical reactions. amu, atomic mass units.

barrier. The relatively high evaporation flux in the first size regime indicates the existence of a nucleation barrier and that the critical cluster in the atmospheric nucleation point of view is in the second size regime at 1.5 ± 0.3 nm. In any case, we observe a starting point of a phase transition, which includes chemical reactions to form condensable vapors, heteromolecular clustering of vapor molecules, and the subsequent growth of clusters to 2 and 3 nm.

In summary, we can identify three separate size regimes characterized by different mobility diameter ranges for neutral clusters (Fig. 4). Our findings demonstrate an observational-based framework on atmospheric aerosol formation that, in a consistent way, combines (i) molecules, small atmospheric clusters, and growing nanoparticles; (ii) sulfuric acid, strong bases, and organic vapors; and (iii) various dynamical processes. This framework confirms that atmospheric aerosol formation is essentially a two-step process, as suggested based on theoretical arguments (*21*, *28*) and some laboratory experiments (*29*). In the first step—in the second size regime—atmospheric nucleation or the formation of stabilized clusters will occur. The second step, characterized by enhanced cluster growth rates due to the activation of the growing clusters by organic vapors, is initiated in the third size regime just below 2 nm. This second step determines the formation rate of 3-nm particles and is efficient only during periods of active aerosol formation. Our findings emphasize the important role of organic compounds in atmospheric aerosol formation; in the radiative forcing that results when these particles grow to larger sizes; and in the associated feedbacks involving the biosphere, clouds, and climate (*30*, *31*).

References and Notes

1. M. Kulmala *et al.*, *J. Aerosol Sci.* **35**, 143 (2004).
2. R. Zhang, A. Khalizov, L. Wang, M. Hu, W. Xu, *Chem. Rev.* **112**, 1957 (2012).
3. D. V. Spracklen *et al.*, *Atmos. Chem. Phys.* **6**, 5631 (2006).
4. J. Merikanto, D. V. Spracklen, G. W. Mann, S. J. Pickering, K. S. Carslaw, *Atmos. Chem. Phys.* **9**, 8601 (2009).
5. J. Kazil *et al.*, *Atmos. Chem. Phys.* **10**, 10733 (2010).
6. R. Makkonen *et al.*, *Atmos. Chem. Phys.* **12**, 1515 (2012).
7. P. Hari, M. Kulmala, *Boreal Env. Res.* **10**, 315 (2005).
8. M. Dal Maso *et al.*, *Boreal Env. Res.* **10**, 323 (2005).
9. A. Hirsikko *et al.*, *Atmos. Chem. Phys.* **11**, 767 (2011).
10. R. Weber *et al.*, *Chem. Eng. Commun.* **151**, 53 (1996).
11. V.-M. Kerminen *et al.*, *Atmos. Chem. Phys.* **10**, 10829 (2010).
12. M. Sipilä *et al.*, *Science* **327**, 1243 (2010).
13. C. Kuang *et al.*, *Atmos. Chem. Phys.* **12**, 3573 (2012).
14. T. Kurtén, V. Loukonen, H. Vehkamäki, M. Kulmala, *Atmos. Chem. Phys.* **8**, 4095 (2008).
15. V. Loukonen *et al.*, *Atmos. Chem. Phys.* **10**, 4961 (2010).
16. I. K. Ortega *et al.*, *Atmos. Chem. Phys.* **12**, 225 (2012).
17. P. Paasonen *et al.*, *Atmos. Chem. Phys.* **12**, 9113 (2012).
18. J. Zhao *et al.*, *Atmos. Chem. Phys.* **11**, 10823 (2011).
19. T. Petäjä *et al.*, *Phys. Rev. Lett.* **106**, 228302 (2011).
20. B. R. Bzdek, D. P. Ridge, M. V. Johnston, *Atmos. Chem. Phys.* **11**, 8735 (2011).
21. M. Kulmala, V.-M. Kerminen, T. Anttila, A. Laaksonen, C. D. O'Dowd, *J. Geophys. Res.* **109**, D04205 (2004).
22. A. Metzger *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **107**, 6646 (2010).
23. P. Paasonen *et al.*, *Atmos. Chem. Phys.* **10**, 11223 (2010).
24. L. Wang *et al.*, *Nat. Geosci.* **3**, 238 (2010).
25. N. M. Donahue, E. R. Trump, J. R. Pierce, I. Riipinen, *Geophys. Res. Lett.* **38**, L16801 (2011).
26. M. Ehn *et al.*, *Atmos. Chem. Phys.* **12**, 5113 (2012).
27. M. Kulmala *et al.*, *Science* **318**, 89 (2007).
28. M. Kulmala, L. Pirjola, J. M. Mäkelä, *Nature* **404**, 66 (2000).
29. P. E. Wagner, R. Strey, *J. Phys. Chem. B* **105**, 11656 (2001).
30. M. Kulmala *et al.*, *Atmos. Chem. Phys.* **4**, 557 (2004).
31. K. Carslaw *et al.*, *Atmos. Chem. Phys.* **10**, 1701 (2010).

Acknowledgments: This work was funded by the European Research Council Project (grant nos. 227463-ATMNUCLE and 257360-MOCAPAF). We thank the Academy of Finland Centre of Excellence program and related projects (grants nos. 1118615 and 1127372) for support.

Supplementary Materials

www.sciencemag.org/cgi/content/full/339/6122/943/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S24
Tables S1 to S7
References (*32*–*65*)

12 July 2012; accepted 12 December 2012
10.1126/science.1227385

Fruit Flies Medicate Offspring After Seeing Parasites

Balint Z. Kacsoh, Zachary R. Lynch, Nathan T. Mortimer, Todd A. Schlenke*

Hosts have numerous defenses against parasites, of which behavioral immune responses are an important but underappreciated component. Here we describe a behavioral immune response that *Drosophila melanogaster* uses against endoparasitoid wasps. We found that when flies see wasps, they switch to laying eggs in alcohol-laden food sources that protect hatched larvae from infection. This change in oviposition behavior, mediated by neuropeptide F, is retained long after wasps are removed. Flies respond to diverse female larval endoparasitoids but not to males or pupal endoparasitoids, showing that they maintain specific wasp search images. Furthermore, the response evolved multiple times across the genus *Drosophila*. Our data reveal a behavioral immune response based on anticipatory medication of offspring and outline a nonassociative memory paradigm based on innate parasite recognition by the host.

Although immune systems are often thought of as a set of infection-responsive molecules and cells within a host, they comprise a much more diverse array of biological structures and processes that collectively protect an organism from infection. Medication, the prophylactic (pre-infection) or therapeutic (post-infection) use of substances found in the environment to combat infection, is a type of behavioral immune mechanism (*1*). Medication requires the recognition of infection, or infection risk, by the host, leading to the use of a substance directed against the identified parasite (*2*, *3*). Endoparasitoid wasps are a serious threat to flies in nature (*4*), and we recently showed that infected *Drosophila melanogaster* larvae preferentially consume toxic levels of alcohol because the benefit of alcohol-mediated wasp death outweighs the cost to flies of alcohol consumption, an example of therapeutic self-medication (*5*). In the present study, we tested whether adult fruit flies choose to lay their eggs in food containing toxic levels of alcohol when wasps are present in the environment as a means of prophylactically medicating their offspring against infection.

We tested the oviposition preferences of adult female *D. melanogaster* by placing 300 flies in population cages containing two food dishes, one of which contained 6% ethanol by volume (Fig. 1A). Flies were housed with or without 50 female wasps, and fly eggs were counted from separate sets of dishes 24 and 48 hours later. Control flies preferred to oviposit on dishes containing no ethanol; but in the presence of female *Leptopilina heterotoma*, a common wasp parasite of *D. melanogaster* larvae in nature (*6*), flies laid a significantly greater proportion of eggs on ethanol dishes at both time points (Fig. 1B and table S1). The flies displayed no such alcohol preference in the presence of male wasps. To determine the extent of fly preference for alcohol-laden oviposition sites in the presence of female wasps, flies were given a choice of various concentrations of ethanol. Control flies preferred to oviposit on dishes containing 3% ethanol (Fig. 1C and table S2), which is consistent with the known benefits to fly larvae of low-level alcohol consumption and the costs of higher-level consumption (*5*, *7*–*9*). In the presence of wasps, however, flies overwhelmingly preferred to oviposit on dishes containing ethanol concentrations corresponding to the highest levels found in nature (12 and 15%) (Fig. 1C and table S2) (*10*).

To determine whether the fly oviposition switch is adaptive, we measured offspring eclosion success in various oviposition setups. In the absence of wasps, the offspring of flies in cages with only 0% ethanol dishes had significantly higher eclosion success than offspring from flies given 6% alcohol food, demonstrating that there is normally a fitness detriment to ovipositing in food with such high alcohol levels (Fig. 1D). When female wasps were present, however, the offspring of flies given an opportunity to oviposit on alcohol-laden food had significantly higher eclosion success than offspring of flies given no such opportunity (Fig. 1E). This prophylaxis probably arises from both decreased offspring infection and increased offspring success at curing infections (*5*). Such an induced fly behavioral immune response may serve as an alternative to the energetically costly cellular encapsulation response that flies mount against wasp eggs.

Mutant strains were used to determine whether flies require olfactory or visual cues to sense wasps. *Or83b* mutants fail to respond to most olfactory stimuli (*11*) but retained an oviposition preference for ethanol food in the presence of wasps (Fig. 2A, fig. S1A, and table S3), suggesting that this general olfactory receptor is not

Department of Biology, Emory University, 1510 Clifton Road NE, Atlanta, GA 30322, USA.

*To whom correspondence should be addressed. E-mail: tschlen@emory.edu

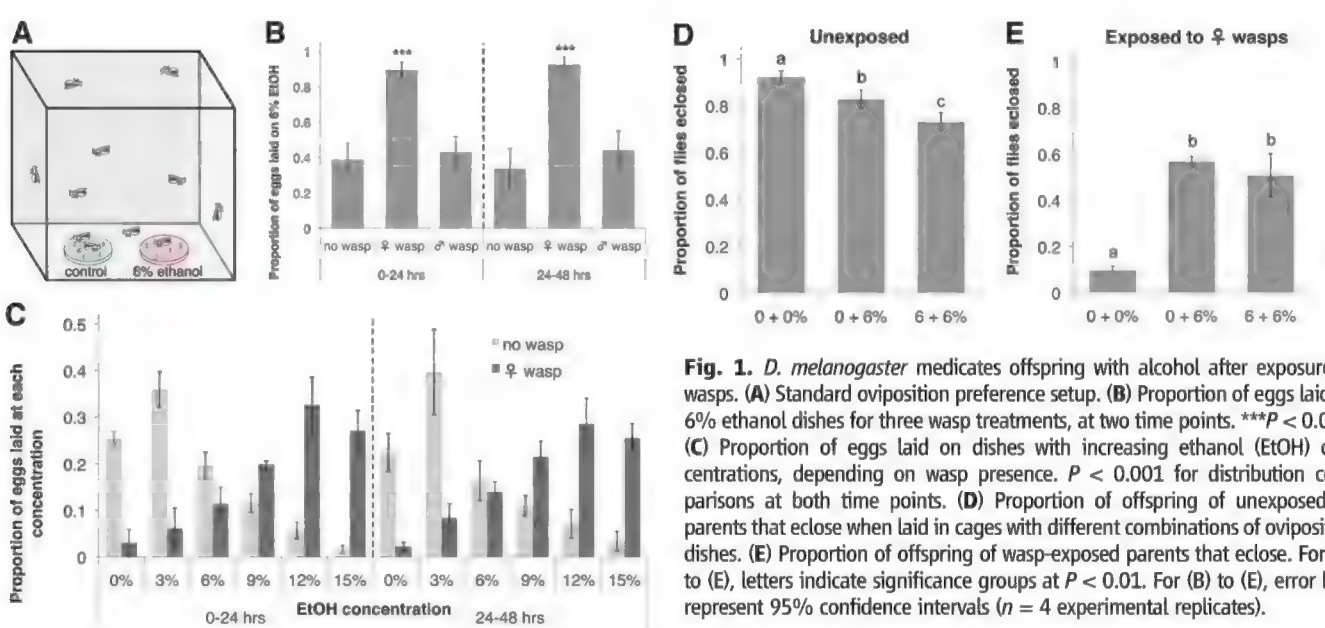
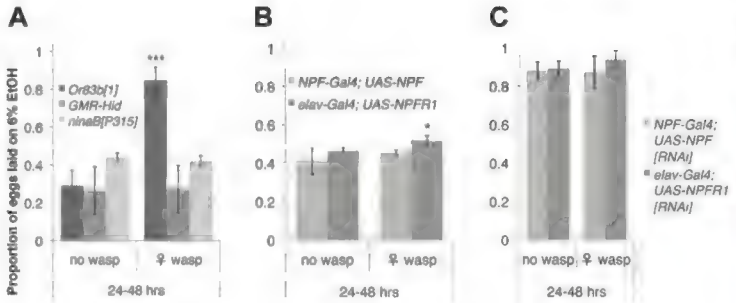


Fig. 1. *D. melanogaster* medicates offspring with alcohol after exposure to wasps. (**A**) Standard oviposition preference setup. (**B**) Proportion of eggs laid on 6% ethanol dishes for three wasp treatments, at two time points. ***$P < 0.001$. (**C**) Proportion of eggs laid on dishes with increasing ethanol (EtOH) concentrations, depending on wasp presence. $P < 0.001$ for distribution comparisons at both time points. (**D**) Proportion of offspring of unexposed fly parents that eclose when laid in cages with different combinations of oviposition dishes. (**E**) Proportion of offspring of wasp-exposed parents that eclose. For (D) to (E), letters indicate significance groups at $P < 0.01$. For (B) to (E), error bars represent 95% confidence intervals ($n = 4$ experimental replicates).



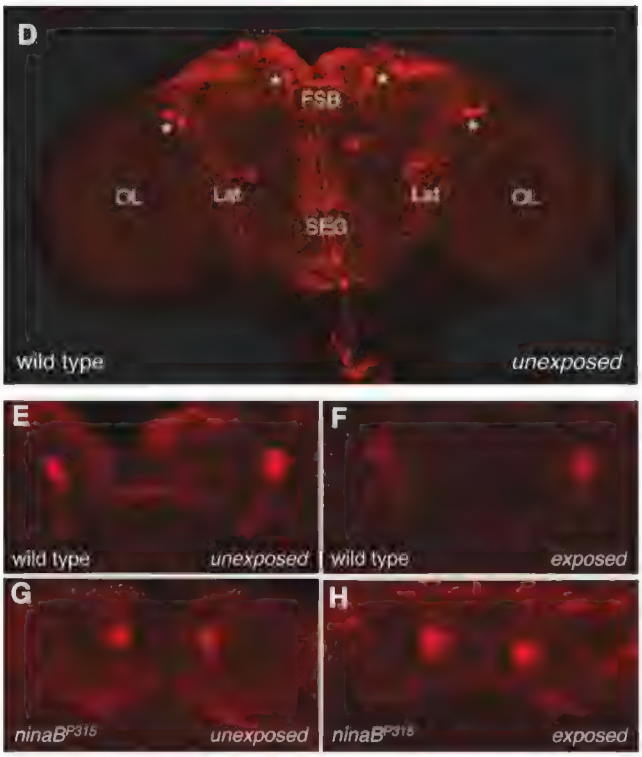


Fig. 2. Sight and NPF signaling control fly ability to sense and respond to wasps. (**A** to **C**) Proportion of eggs laid on ethanol dishes by (A) smell and sight mutants, (B) *NPF* and *NPFR1* overexpression mutants, and (C) *NPF* and *NPFR1* knockdown mutants in the presence and absence of wasps. For (A) to (C), the *y* axis is the same; error bars represent 95% confidence intervals (*$P < 0.05$, ***$P < 0.001$, $n = 4$). (**D**) NPF immunostain of an unexposed fly brain. *, NPF-expressing neurons; FSB, fan-shaped body; Lat, lateral regions; SEG, subesophageal ganglion; OL, optic lobes. (**E** to **H**) NPF-immunostained fan-shaped bodies from control and sight mutant flies unexposed or exposed to wasps.

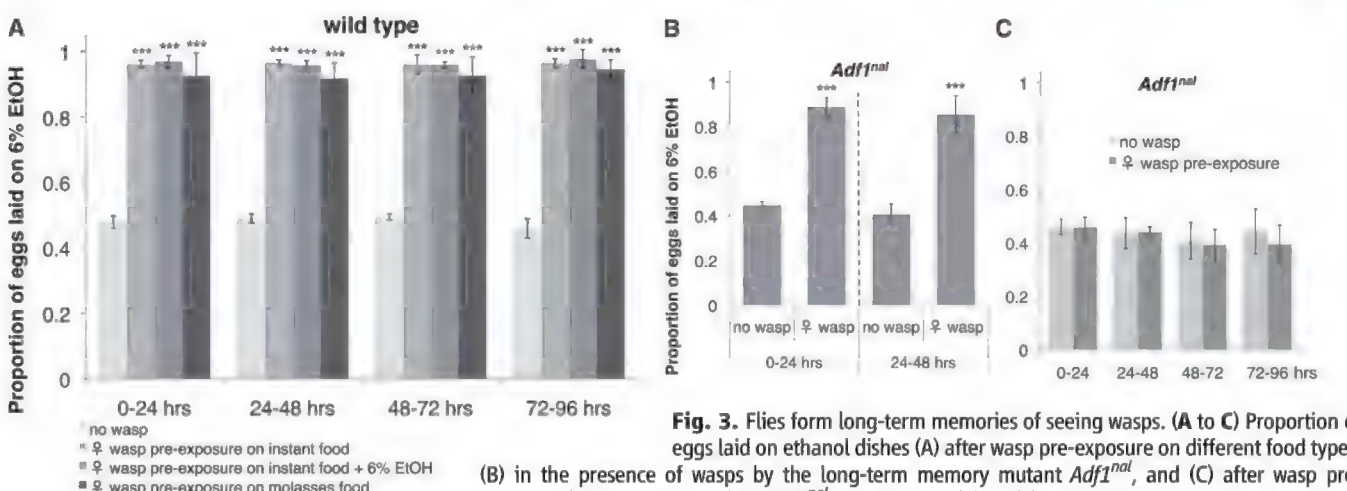


Fig. 3. Flies form long-term memories of seeing wasps. (**A** to **C**) Proportion of eggs laid on ethanol dishes (A) after wasp pre-exposure on different food types, (B) in the presence of wasps by the long-term memory mutant *Adf1*^{nal}, and (C) after wasp pre-exposure (on molasses food) in *Adf1*^{nal} mutants. For (A) to (C), all *y* axes are the same; error bars represent 95% confidence intervals (*$P < 0.05$, ***$P < 0.001$, $n = 4$).

required for wasp detection or alcohol sensing (*12*). *GMR-Hid* flies express an apoptotic activator in the developing retina leading to dramatically reduced eyes (*13*), and *ninaB*^{P315} mutants fail to synthesize rhodopsin, eliminating vision while leaving the basic morphology of the eye intact (*14*). Neither vision mutant showed an oviposition preference for ethanol food in the presence of wasps (Fig. 2A, fig. S1A, and table S3), indicating that flies rely on sight to sense wasps in their environment and initiate the oviposition preference switch.

Reduced expression of neuropeptide F (NPF) and its receptor (NPFR1) in fly brains increases ethanol tolerance and preference, similarly to reducing NPY in mammals (*15*, *16*). Using the yeast *Gal4-UAS* transcription factor–promoter system, we found that *NPF-Gal4*–driven ectopic expression of *NPF* eliminated the wasp-induced ethanol oviposition preference (Fig. 2B; fig. S1, B, and D to G; and table S3). Flies with *elav-Gal4*–driven pan-neuronal ectopic expression of *NPFR1* showed a weak but significant increase in alcohol preference in the presence of wasps, but there was a dramatic reduction in oviposition preference for alcohol food as compared to wild-type flies (Fig. 1B). RNA interference–mediated knockdown of *NPF* and *NPFR1* levels in the brain, using *NPF-Gal4* and *elav-Gal4* drivers, respectively, led to increased ethanol oviposition preference regardless of wasp presence (Fig. 2C; fig. S1, C to E and H to I; and table S3).

These results suggest that the visual perception of wasps by flies might cause decreased NPF levels in fly brains. We immunostained fly brains with NPF antiserum after exposure to wasps (Fig. 2, D to H) and found a marginally significant decrease in whole-brain fluorescence levels in wild-type flies exposed to wasps for 24 hours but no such change in sight-impaired *ninaB*^{P315} flies (fig. S1J). *NPF*-expressing neurons send projections to the fan-shaped body, subesophageal ganglion, and lateral regions of the lower central brain (Fig. 2D) (*16*, *17*). There was a significant loss of NPF immunofluorescence in the fan-shaped body of wild-type flies exposed to wasps, whereas the other NPF-positive regions remained unchanged (Fig. 2, E to F, and fig. S1, K to O). Despite constitutive

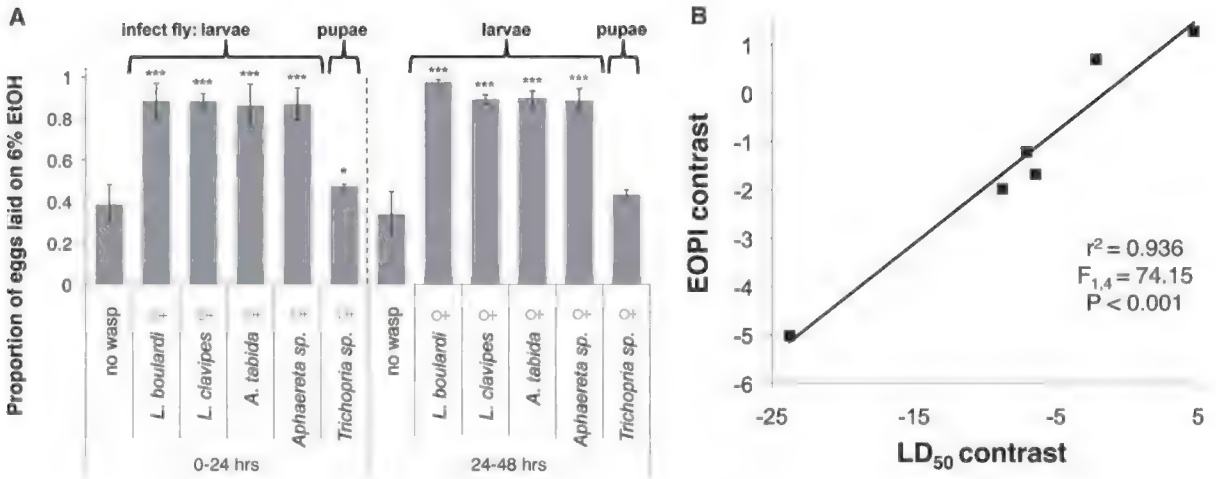


Fig. 4. Breadth of fly search images and the evolution of medication behavior. (**A**) Proportion of eggs laid on ethanol dishes in response to different wasp species. Error bars represent 95% confidence intervals (*$P < 0.05$, ***$P < 0.001$, $n = 4$). (**B**) Correlation between phylogenetically independent contrasts for ethanol tolerance (LD_{50}) and ethanol oviposition preference index (EOPI) across seven *Drosophila* species.

differences between wild-type and $ninaB^{P315}$ flies, there was no difference in the staining of the fan-shaped body or other brain regions between unexposed and wasp-exposed $ninaB^{P315}$ flies (Fig. 2, G to H, and fig. S1, K to O). The fan-shaped body is part of the central complex of the fly brain that regulates both visual pattern recognition and ethanol-stimulated locomotion (*18*, *19*).

To determine how long fly oviposition preferences are altered after sensing wasps (*20*), we "pre-exposed" flies to wasps before assaying oviposition preferences in cages devoid of wasps. Flies pre-exposed to wasps showed a strong preference for the ethanol oviposition dishes across a 4-day choice assay (Fig. 3A and table S4). The oviposition switch occurred even when flies were pre-exposed to wasps in food bottles containing alcohol or completely different media, demonstrating that flies do not learn to associate wasp presence with food type (Fig. 3A: $P > 0.120$ for preference comparisons between pre-exposure food types at each time point, table S4). Thus, the oviposition preference switch that flies maintain after seeing wasps is a natural example of nonassociative memory (*21*).

The *D. melanogaster* transcription factor Adf1 is required for long-term memory formation and for regulating *Alcohol dehydrogenase* expression (*21*, *22*). $Adf1^{nal}$, a mutant that has normal early memory but lacks long-term memory, switched its oviposition preference to alcohol food in the presence of wasps like wild-type flies (Fig. 3B and table S4). When $Adf1^{nal}$ mutants were pre-exposed to wasps and then put into cages without wasps, however, the flies failed to retain the wasp exposure memory, showing no oviposition preference for alcohol food (Fig. 3C and table S4). This result is not explained by any reduced ethanol tolerance of $Adf1^{nal}$ flies (fig. S2, A and B). Furthermore, both vision and NPF signaling were required for initiating memory formation (fig. S2, C to F, and table S4). These data suggest that a single protein (Adf1) may simultaneously be responsible for memory of wasp presence and regulation of a gene that controls tolerance to the alcohol-laden food flies become permanently attracted to, and they also outline a simple model of wasp-mediated alcohol seeking in flies (fig. S3).

Innate and learned search images are important for numerous organismal interactions (*23*, *24*). To further delimit the innate wasp search images *D. melanogaster* maintains, we assayed fly oviposition behavior during exposure to two more Figitid wasps, *L. boulardi* and *L. clavipes*; two Braconids, *Aphaereta* sp. and *Asobara tabida*; and a Diapriid pupal endoparasitoid, *Trichopria* sp. (*25*). Flies preferred ethanol-laden oviposition dishes in the presence of females of each endoparasitoid species that infects fly larvae, but the preference was weaker upon exposure to the wasp that infects fly pupae, reaching only marginal significance at 0 to 24 hours (Fig. 4A and table S1). This might make adaptive sense given that *Drosophila* larvae often move off their food source before pupating. Thus, *D. melanogaster* has evolved search images specific enough to distinguish female from male *L. heterotoma* and a pupal endoparasitoid from larval endoparasitoids, but broad enough to recognize two families of larval endoparasitoids (see fig. S4 for wasp images).

To understand the evolution of this behavioral immune mechanism, we tested oviposition preferences in six more *Drosophila* species and found that three species (*D. simulans*, *D. hydei*, and *D. virilis*) showed a significant increase in oviposition preference for alcohol dishes in the presence of female *L. heterotoma* wasps, whereas the other species actively avoided ethanol dishes regardless of wasp presence (fig. S5, A to F, and table S5). Ethanol tolerance median lethal dose (LD_{50}) values for each fly species (fig. S5, G to L) were positively correlated with the ethanol oviposition preference index, a measure of how strongly the preference for ovipositing in ethanol food increases in response to wasp exposure (fig. S5M). This relation was even stronger when the phylogenetic relationships between the *Drosophila* species were taken into account, using phylogenetically independent contrasts (*26*) (Fig. 4B). Thus, alcohol tolerance and the alcohol oviposition preference switch have coevolved across the genus *Drosophila*.

References and Notes

1. J. C. de Roode, T. Lefevre, *Insects* **3**, 789 (2012).
2. T. Lefèvre, L. Oliver, M. D. Hunter, J. C. De Roode, *Ecol. Lett.* **13**, 1485 (2010).
3. M. S. Singer, K. C. Mace, E. A. Bernays, *PLoS ONE* **4**, e4796 (2009).
4. F. Fleury *et al.*, *Genetica* **120**, 181 (2004).
5. N. F. Milan, B. Z. Kacsoh, T. A. Schlenke, *Curr. Biol.* **22**, 488 (2012).
6. F. Fleury, P. Gibert, N. Ris, R. Allemand, *Adv. Parasitol.* **70**, 3 (2009).
7. S. S. Chawla, J. M. Perron, C. Radouco-Thomas, *Can. Entomol.* **113**, 315 (1981).
8. S. W. McKechnie, B. W. Geer, *Insect Biochem.* **14**, 231 (1984).
9. P. A. Parsons, S. M. Stanley, G. E. Spence, *Aust. J. Zool.* **27**, 747 (1979).
10. J. B. Gibson, T. W. May, A. V. Wilks, *Oecologia* **51**, 191 (1981).
11. L. B. Vosshall, H. Amrein, P. S. Morozov, A. Rzhetsky, R. Axel, *Cell* **96**, 725 (1999).
12. E. C. Kong *et al.*, *PLoS ONE* **5**, e9954 (2010).
13. L. Goyal, K. McCall, J. Agapite, E. Hartwieg, H. Steller, *EMBO J.* **19**, 589 (2000).
14. J. von Lintig, A. Dreher, C. Kiefer, M. F. Wernet, K. Vogt, *Proc. Natl. Acad. Sci. U.S.A.* **98**, 1130 (2001).
15. T. E. Thiele, D. J. Marsh, L. Ste Marie, I. L. Bernstein, R. D. Palmiter, *Nature* **396**, 366 (1998).
16. T. Q. Wen, C. A. Parrish, D. Xu, Q. Wu, P. Shen, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 2141 (2005).
17. M. R. Brown *et al.*, *Peptides* **20**, 1035 (1999).
18. G. Liu *et al.*, *Nature* **439**, 551 (2006).
19. R. Strauss, *Curr. Opin. Neurobiol.* **12**, 633 (2002).
20. T. Lefèvre, J. C. de Roode, B. Z. Kacsoh, T. A. Schlenke, *Biol. Lett.* **8**, 230 (2012).
21. C. Margulies, T. Tully, J. Dubnau, *Curr. Biol.* **15**, R700 (2005).
22. B. P. England, U. Heberlein, R. Tjian, *J. Biol. Chem.* **265**, 5086 (1990).

23. R. Menzel, in *Experimental Behavioral Ecology and Sociobiology*, B. Holldobler, M. Lindauer, Eds. (Sinauer Associates, Sunderland, MA, 1985), pp. 55–74.
24. A. T. Pietrewicz, A. C. Kamil, *Science* **204**, 1332 (1979).
25. B. Z. Kacsoh, T. A. Schlenke, *PLoS ONE* **7**, e34721 (2012).
26. P. H. Harvey, A. Purvis, *Nature* **351**, 619 (1991).

Acknowledgments: We thank P. Shen for NPF antiserum; P. Shen, K. Moberg, and S. Sanyal for *Drosophila* strains; and J. van Alphen and B. Wertheim for wasp strains. This work was supported by NIH grant AI081879 to T.A.S. and the Integrated Cellular Imaging Microscopy Core of the Emory Neuroscience NINDS Core Facilities grant, P30NS055077. Data are deposited in the Dryad Repository: http://dx.doi.org/10.5061/dryad.j5g7m.

Supplementary Materials
www.sciencemag.org/cgi/content/full/339/6122/947/DC1
Materials and Methods
Figs. S1 to S5
Tables S1 to S5

3 September 2012; accepted 10 January 2013
10.1126/science.1229625

Precise Maps of RNA Polymerase Reveal How Promoters Direct Initiation and Pausing

Hojoong Kwak, Nicholas J. Fuda, Leighton J. Core,* John T. Lis*

Transcription regulation occurs frequently through promoter-associated pausing of RNA polymerase II (Pol II). We developed a precision nuclear run-on and sequencing (PRO-seq) assay to map the genome-wide distribution of transcriptionally engaged Pol II at base pair resolution. Pol II accumulates immediately downstream of promoters, at intron-exon junctions that are efficiently used for splicing, and over 3′ polyadenylation sites. Focused analyses of promoters reveal that pausing is not fixed relative to initiation sites, nor is it specified directly by the position of a particular core promoter element or the first nucleosome. Core promoter elements function beyond initiation, and when optimally positioned they act collectively to dictate the position and strength of pausing. This "complex interaction" model was tested with insertional mutagenesis of the *Drosophila Hsp70* core promoter.

Tracking the accumulation of RNA polymerase II (Pol II) along genes reveals potential points of regulation (*1*). For example, a rate-limiting step in early elongation, known as promoter-proximal pausing, has revealed a major regulatory block in the transition to productive elongation in *Drosophila* and mammals (*2*–*8*). Also, less extensive but substantial accumulation of Pol II over the 3′ cleavage/polyadenylation region of genes is proposed to facilitate 3′ processing and transcription termination (*9*, *10*). Finally, the interplay of transcription rate and splicing efficiency (*11*) might be reflected in the selective accumulation of Pol II at splice junctions.

Promoter-associated Pol II pausing is a culmination of intrinsic interactions between Pol II and the underlying DNA, as well as extrinsic stabilization by protein complexes (*12*). Protein factors such as NELF (negative elongation factor) and DSIF [5,6-dichloro-1-β-D-ribofuranosylbenzimidazole (DRB) sensitivity-inducing factor] (*3*, *13*), DNA elements (*14*, *15*), DNA sequence composition (*16*), nascent RNA processing (*16*), and nucleosomes (*17*) can influence pausing. Understanding how these elements and factors function mechanistically requires a high-resolution view of their spatial relationship. Current tools for precise tracking of the location and status of Pol II in vivo have distinct limitations (*18*). Chromatin immunoprecipitation–based methods that collect Pol II or associated RNAs do not distinguish paused Pol II from other Pol II–RNA complexes (*16*, *18*, *19*). The genome-wide nuclear run-on (GRO-seq) approach (*6*–*8*) circumvents these issues by high specific enrichment of nascent transcripts associated with actively engaged

Department of Molecular Biology and Genetics, Cornell University, Ithaca, NY 14853, USA.

*To whom correspondence should be addressed. E-mail: ljc37@cornell.edu (L.J.C.); jtl10@cornell.edu (J.T.L.)

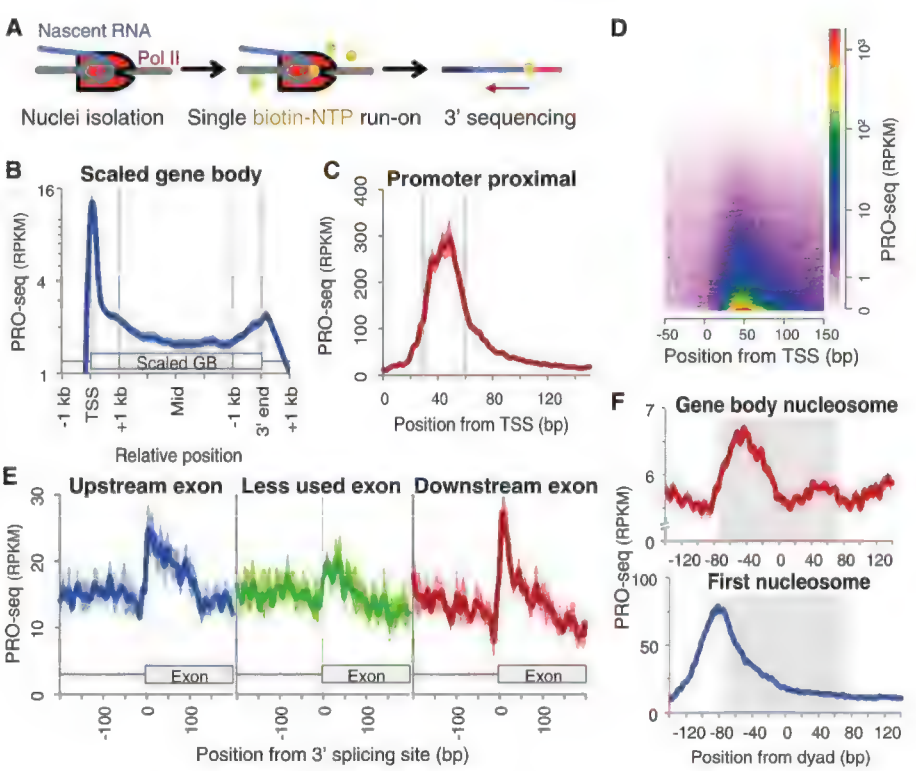


Fig. 1. Accumulation of Pol II at promoters, 3′ ends, 3′ splice sites, and nucleosomes. (**A**) Schematic of PRO-seq. (**B**) Average PRO-seq profile of nonoverlapping genes ($n = 6309$) for the sense strand. Gene body (GB) regions (+1 kb from the 5′ end to −1 kb from the 3′ end) are scaled to 4 kb. Read counts are adjusted to RPKM (reads per kilobase per million mapped reads). Shaded margins surrounding the average plot represent SEM. (**C**) High-resolution PRO-seq profile from the TSS to +150 bp ($n = 16{,}746$). (**D**) Heat map visualization of PRO-seq profile of the annotated genes. Genes are arranged by their increasing PRO-seq density. (**E**) Average PRO-seq profile at 3′ splicing sites of less-used exons and their flanking exons ($n = 242$ each). Less-used exons have RNA-seq densities less than 5% of their flanking exons (fig. S5B). (**F**) Average PRO-seq profile relative to the dyad centers of gene body regions and first nucleosomes. The region occupied by nucleosome is shaded gray.

polymerase, but it has a resolution of only 30 to 50 bases (*18*).

We developed a genome-wide, nuclear run-on assay called PRO-seq that has the sensitivity of GRO-seq but maps Pol II with base pair resolution. PRO-seq uses biotin-labeled ribonucleotide triphosphate analogs (biotin-NTP) for nuclear run-on reactions, allowing the efficient affinity purification of nascent RNAs for high-throughput sequencing from their 3′ ends (Fig. 1A and fig. S1A). Supplying only one of the four biotin-NTPs (adenosine, cytidine, guanosine, or uridine triphosphate) restricts Pol II to incorporate a single or at most a few identical bases, resulting in sequence reads that have the same 3′ end base within each library (table S1). Moreover, the incorporation of the first biotin base inhibits further transcript elongation, ensuring base pair resolution (fig. S2).

The average profile of PRO-seq density (Fig. 1B) revealed pausing of Pol II immediately downstream of the transcription start site (TSS) (Fig. 1, C and D) and accumulation of Pol II at 3′ cleavage/polyadenylation sites, consistent with previous GRO-seq studies (fig. S3) (*20*, *21*). Interestingly, Pol II also accumulated near 3′ splicing sites at spliced exons, but less often at skipped exons (Fig. 1E and fig. S4), which suggests that splicing decisions are connected to differential rates of Pol II elongation through splice junctions (*11*). Although we have insufficient sequencing coverage to quantify Pol II accumulation at particular 3′ splice sites, our composite analyses support a functional coupling between elongation and splicing.

The highest density of PRO-seq reads mapped within positions +30 to +60 from the TSS (Fig. 1, C and D), providing a higher-resolution view of paused Pol II as mapped by GRO-seq (*21*). Moreover, the pattern of pausing by PRO-seq is consistent with the positions and levels of short nuclear-capped RNAs (scRNAs; fig. S3) (*16*). Additionally, we found that PRO-seq maps correspond precisely to positions of engaged Pol II observed in intact cells seen by previous permanganate footprints of transcription bubbles (fig. S2G).

Nucleosomes are known to act as barriers to Pol II (*12*). In the bodies of genes, the average PRO-seq density showed a relative increase around position −40 from the previously mapped (*22*) nucleosome centers (Fig. 1F and fig. S5A). This is consistent with measurements of strong DNA-nucleosome interactions (*23*) and measured impediments to Pol II transcription through nucleosomes measured in vitro and in yeast (*20*, *24*). However, the PRO-seq density relative to the first (+1) nucleosome was different (Fig. 1F), with the average PRO-seq density at a maximum around position −80 from the nucleosome centers. Thus, the bulk of promoter-proximal pausing is inconsistent with a standard nucleosome barrier model, at least for *Drosophila*, and is more consistent with tethering of polymerases near the promoter (*21*).

Although the average promoter-associated pause location is around position +40 from the TSS, pausing is far from uniform. Some genes have more proximal and focused pausing, whereas others have distal and dispersed pausing (Fig. 2A). We systematically assessed genome-wide pausing positions relative to the TSS and their dispersion to identify two characteristic groups of promoters: focused proximal (*Prox*) and dispersed distal (*Dist*) promoters (Fig. 2B and fig. S6). The *Prox* and *Dist* pausing patterns could arise from a fixed length of elongation from initiation sites that have the same dispersion or from variable lengths of elongation from more focused initiation sites. Distinguishing between these possibilities requires precise mapping of the initiation sites, using the same pool of Pol II–engaged nascent RNAs. Therefore, we modified the PRO-seq method to detect initiation sites (PRO-cap; fig. S1B) and compared the degree of variation in the initiation and pause sites. We observed that both *Prox* and *Dist* genes have relatively focused initiation in general (Fig. 2, A and D, and fig. S2G) and that pausing is overall more dispersed than initiation (fig. S7C). Nonetheless, the degree of focused initiation—the fraction of initiation arising at a single TSS—is higher for *Prox* genes, and genes with more focused initiation also have more proximal pausing (Fig. 2D). These findings indicate that although pausing is not fixed to initiation, the mechanisms that produce focused initiation affect the resulting pattern of pausing.

In an effort to otherwise explain the differential patterns of pausing, we first compared the nucleosome occupancy around *Prox* and *Dist* promoters. *Prox* promoters have less nucleosome occupancy than *Dist* promoters (fig. S5B), and some of the Pol II at *Dist* promoters appears to have more intimate contact with the first nucleosome (fig. S5C). These results (and Fig. 1F) support a nucleosome-independent mechanism of pausing for *Prox* promoters, whereas a subset of *Dist* promoters could have a component of pausing

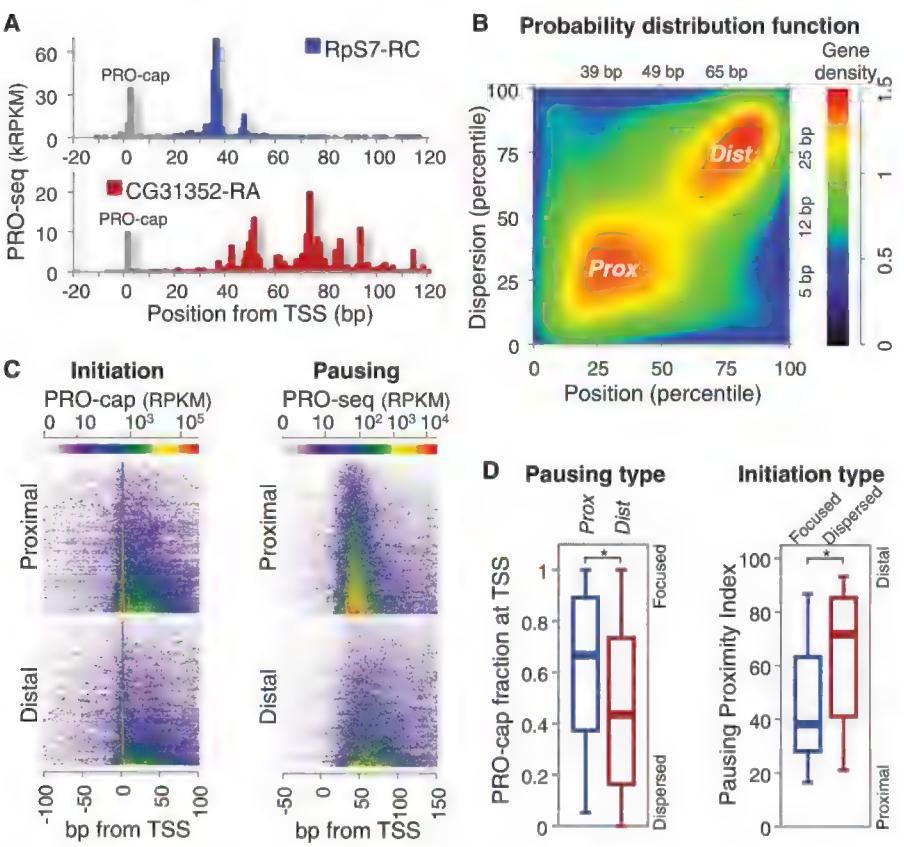


Fig. 2. Variations of the pause sites and TSSs. (**A**) Examples of highly paused genes with different pausing patterns. Initiation sites from PRO-cap mapping are shown in gray. (**B**) Distribution of paused genes (n = 3225) by pausing position and dispersion percentiles. Focused proximal (*Prox*, n = 848) and dispersed distal (*Dist*, n = 846) groups are indicated; axis units in base pairs are also shown. (**C**) Heat map of initiation (PRO-cap) and pausing (PRO-seq) for *Prox* and *Dist* genes. (**D**) Association between initiation and pausing patterns: TSS focusing in *Prox* versus *Dist* genes (left) and pausing proximity in focused versus dispersed initiation genes (right). TSS focusing is measured as the fraction of PRO-cap reads at the TSS (±1 bp) relative to the sum of reads around the TSS (±50 bp) (*22*). Focused and dispersed initiation genes are the quartiles of the paused genes with the highest and the lowest TSS focusing, respectively. The pausing proximity index is defined by the average of pausing position and dispersion percentiles (fig. S6A). Boxes represent 25th, 50th (median), and 75th percentiles; whiskers are 5th and 95th percentiles. *P < 0.001 [Kolmogorov-Smirnov (KS) test].

that is established by direct nucleosome barriers. Because nucleosome position and occupancy do not explain the bulk of Pol II pausing, we investigated the underlying DNA elements around promoters.

Critical DNA sequence elements within the core promoter direct the position, direction, and efficiency of transcription initiation (*25*). These include the TATA box, initiator (Inr), motif 10 element (MTE), downstream promoter element (*25*), and a recently discovered element implicated in pausing, the pause button (PB) (*15*) (fig. S8A). Core promoter elements were more enriched on *Prox* promoters than on *Dist* promoters (Fig. 3A and fig. S8, B to D). Additionally, when

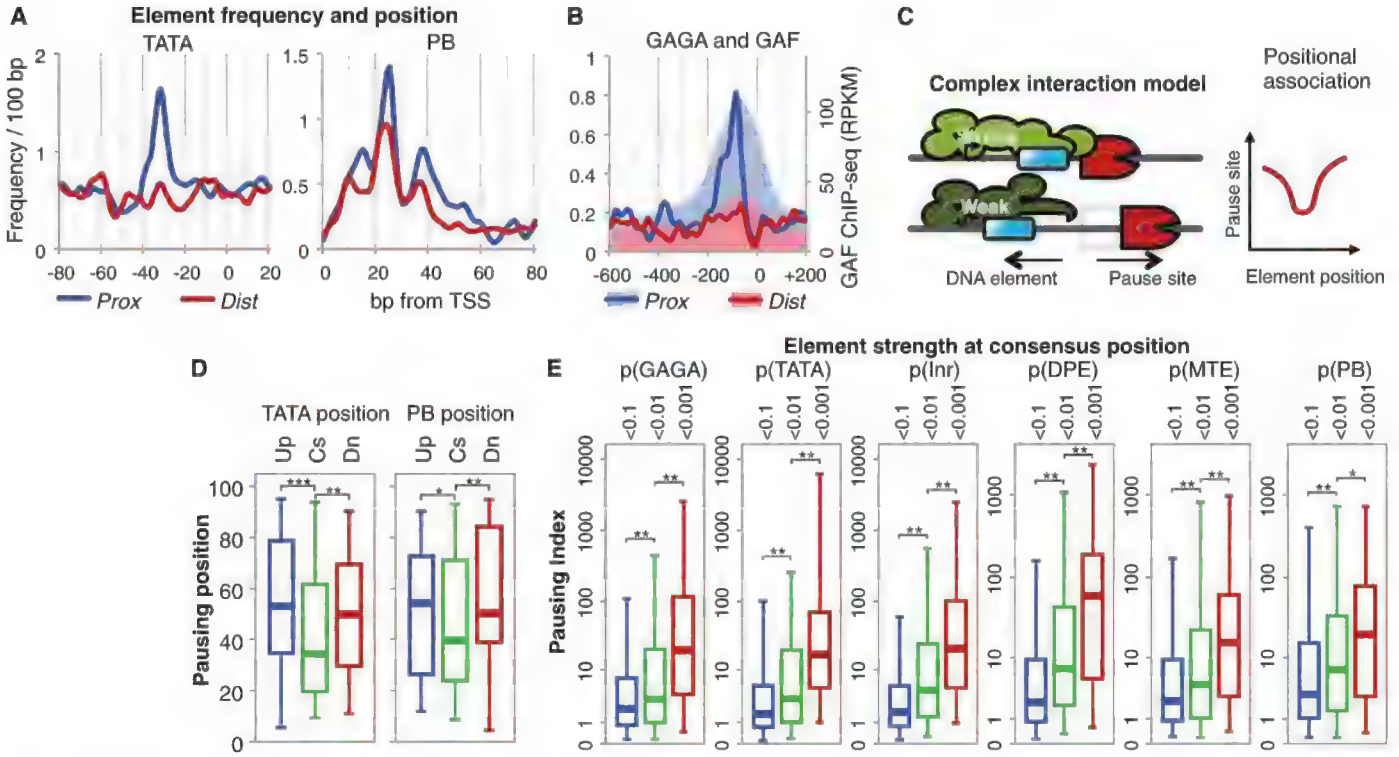


Fig. 3. Relationship between promoter DNA elements and Pol II pausing. (**A**) Frequency of TATA box and PB in *Prox* and *Dist* subsets. The average frequency per gene is shown. (**B**) Frequency of GAGA element (lines) and GAGA-factor binding (shaded areas) (*27*) in *Prox* and *Dist* subsets. (**C**) The complex interaction model of relationship between DNA elements and paused Pol II. The DNA elements (blue) are at their consensus (strong) or slightly upstream (weak) positions; the expected changes of the pausing positions are plotted. (**D**) Pattern of the positional association between DNA elements and pausing positions. Pausing position percentiles are shown for gene subsets according to element position (Up, upstream; Cs, optimal consensus position; Dn, downstream; subset information in table S2). $^{*}P < 0.14$, $^{**}P < 0.06$, $^{***}P < 0.002$ (KS test). (**E**) Association of promoter DNA element strength at consensus positions with pausing index (*6*). Active genes ($n = 5471$) are divided into three subsets according to the distance-weighted *P* values of the DNA elements to the consensus positions (table S3). $^{*}P < 0.01$, $^{**}P < 0.001$ (KS test).

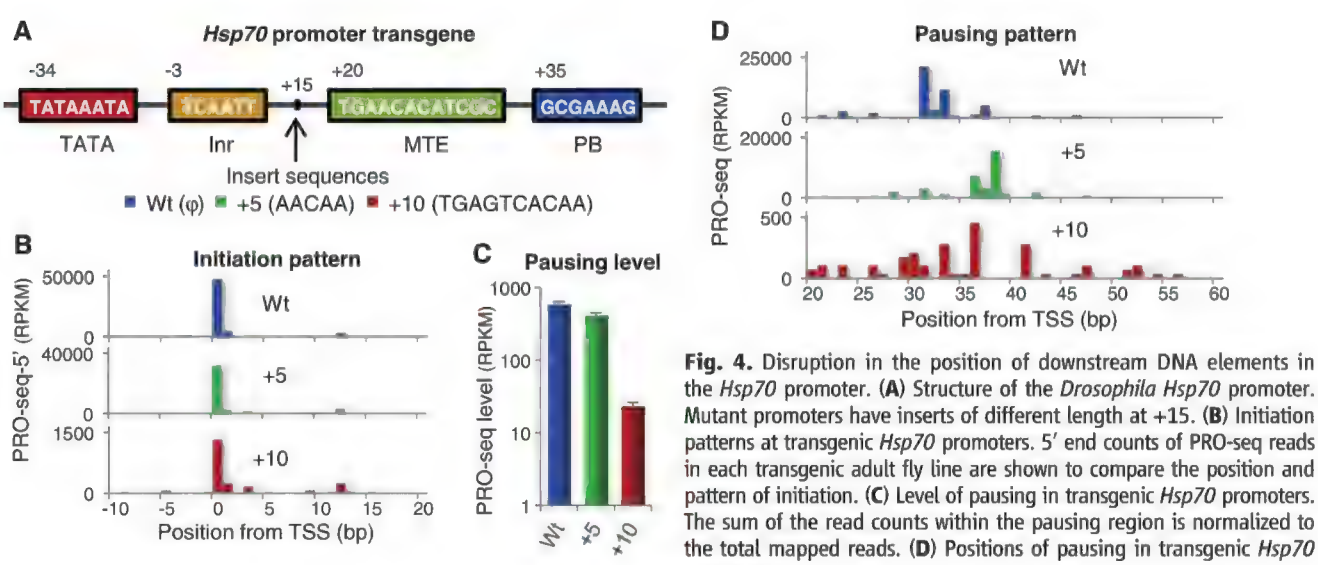


Fig. 4. Disruption in the position of downstream DNA elements in the *Hsp70* promoter. (**A**) Structure of the *Drosophila Hsp70* promoter. Mutant promoters have inserts of different length at +15. (**B**) Initiation patterns at transgenic *Hsp70* promoters. 5′ end counts of PRO-seq reads in each transgenic adult fly line are shown to compare the position and pattern of initiation. (**C**) Level of pausing in transgenic *Hsp70* promoters. The sum of the read counts within the pausing region is normalized to the total mapped reads. (**D**) Positions of pausing in transgenic *Hsp70* promoters.

we searched within the extended promoter regions of *Prox* and *Dist* groups for the presence of 232 additional functional DNA elements (*26*) (fig. S8E), only the GAGA element, residing ~80 base pairs (bp) upstream of the TSS (*3*, *15*), showed strong associations with *Prox* genes (Fig. 3B), as did the level of GAGA-factor binding (*27*). Thus, core promoter elements and GAGA factor appear to play an important role in the mechanism of pausing.

Pausing positions could be determined through direct tethering of elongating Pol II to promoter elements. Alternatively, in a "complex interaction" model, pausing could be mediated through protein complexes that function best when cognate elements are located at specific positions in the core promoter. Thus, if we examine the association of the positions of the DNA elements and the pausing sites in this model, we expect a V-shaped plot of association rather than a simple linear correlation. Displacement of the element from the optimal position will weaken the interactions within the core complex, resulting in downstream scattering and a reduced level of pausing (Fig. 3C). To test this, we examined genes in which a particular promoter DNA element occurs only once, and divided genes into three subsets: the optimal consensus position, upstream, and downstream. Genes with the DNA elements nearest to the consensus positions showed more proximal pausing. Genes with TATA near –30 had more proximal pausing than genes with TATA at positions of –40 or more, showing a V-shaped association (Fig. 3D). This V pattern was observed in both the upstream elements TATA and Inr (fig. S9A) and the downstream elements PB (Fig. 3D) and MTE (fig. S9B). Also, pausing tended to be stronger in genes with the elements at the optimal positions (fig. S9D). Furthermore, the extent of pausing showed strong dependency on the match of the DNA elements to their consensus sequence and consensus positions (Fig. 3E). Together, these patterns of association between core promoter elements and pausing support the complex interaction model and explain the strong and focused pausing on *Prox* promoters.

The complex interaction model depends on both the presence and the correct positioning of core promoter elements. We disrupted the positional relationship of core elements in the well-studied *Drosophila* gene *Hsp70* (*1*). Transgenic fly lines were generated that carry mutant *Hsp70* promoters with spacers inserted at the +15 position between the upstream and downstream promoter elements (Fig. 4A) and were analyzed by PRO-seq (fig. S10). The initiation sites remained constant in these mutant promoters, as indicated by the 5′ ends of the PRO-seq reads (Fig. 4B). The transgenic *Hsp70* without spacers showed a strong pause peak mainly at position +31 (Fig. 4D). When a 5-bp spacer was inserted, the pause peak was shifted 5 to 7 bp downstream from the original site. Because additional bases were transcribed before pausing, the position of pausing is not predetermined by elongation distance. When a 10-bp spacer was inserted, pausing sites became scattered between positions +20 and +60 (Fig. 4D) and had fewer reads (Fig. 4C). Collectively, these results support the core interaction model and suggest that the interaction complex can accommodate a small change (5 bp) in the positional context of the DNA sequences, whereas a larger change (10 bp) results in reduced and dispersed pausing.

The advances in resolution provided by PRO-seq enable the precise and genome-wide assessment of the relationship between promoter-proximal pausing and the core promoter structure. For the strong and tightly clustered pausing of the *Prox* genes, our results provide support for a complex interaction model involving the promoter initiation complex, which can extend up to 30 bp from the TSS (*28*), physically contacting and tethering the pausing complexes. This may share a kinship with bacterial initiation factor σ, which is retained within the early elongation complex and interacts with promoter proximal DNA during transcription pausing in *Escherichia coli* (*29*). It is noteworthy that the *Prox* genes are expressed on average at a lower level but show a broader range of expression (fig. S6D), and that the *Dist* genes are enriched in constitutively active genes (table S5). These results suggest that the mechanistic distinctions have regulatory consequences. A well-structured core promoter may strongly recruit Pol II; however, it can also effectively retain Pol II in a paused configuration close to the TSS, until activation signals allow its escape into productive elongation.

References and Notes

1. N. J. Fuda, M. B. Ardehali, J. T. Lis, *Nature* **461**, 186 (2009).
2. E. B. Rasmussen, J. T. Lis, *Proc. Natl. Acad. Sci. U.S.A.* **90**, 7923 (1993).
3. C. Lee *et al.*, *Mol. Cell. Biol.* **28**, 3290 (2008).
4. G. W. Muse *et al.*, *Nat. Genet.* **39**, 1507 (2007).
5. J. Zeitlinger *et al.*, *Nat. Genet.* **39**, 1512 (2007).
6. L. J. Core, J. J. Waterfall, J. T. Lis, *Science* **322**, 1845 (2008).
7. N. Hah *et al.*, *Cell* **145**, 622 (2011).
8. I. M. Min *et al.*, *Genes Dev.* **25**, 742 (2011).
9. L. J. Core, J. T. Lis, *Science* **319**, 1791 (2008).
10. K. Glover-Cutter, S. Kim, J. Espinosa, D. L. Bentley, *Nat. Struct. Mol. Biol.* **15**, 71 (2008).
11. A. R. Kornblihtt, *Nat. Struct. Mol. Biol.* **13**, 5 (2006).
12. A. Saunders, L. J. Core, J. T. Lis, *Nat. Rev. Mol. Cell Biol.* **7**, 557 (2006).
13. C. H. Wu *et al.*, *Genes Dev.* **17**, 1402 (2003).
14. C. H. Wu *et al.*, *Mol. Cell. Biol.* **21**, 1593 (2001).
15. D. A. Hendrix, J. W. Hong, J. Zeitlinger, D. S. Rokhsar, M. S. Levine, *Proc. Natl. Acad. Sci. U.S.A.* **105**, 7762 (2008).
16. S. Nechaev *et al.*, *Science* **327**, 335 (2010).
17. T. N. Mavrich *et al.*, *Nature* **453**, 358 (2008).
18. S. Nechaev, K. Adelman, *Biochim. Biophys. Acta* **1809**, 34 (2011).
19. L. S. Churchman, J. S. Weissman, *Nature* **469**, 368 (2011).
20. E. Larschan *et al.*, *Nature* **471**, 115 (2011).
21. L. J. Core *et al.*, *Cell Rep.* **2**, 1025 (2012).
22. D. A. Gilchrist *et al.*, *Cell* **143**, 540 (2010).
23. M. A. Hall *et al.*, *Nat. Struct. Mol. Biol.* **16**, 124 (2009).
24. V. A. Bondarenko *et al.*, *Mol. Cell* **24**, 469 (2006).
25. T. Juven-Gershon, J. Y. Hsu, J. W. Theisen, J. T. Kadonaga, *Curr. Opin. Cell Biol.* **20**, 253 (2008).
26. A. Stark *et al.*; Harvard FlyBase curators; Berkeley *Drosophila* Genome Project, *Nature* **450**, 219 (2007).
27. S. Sharma, M. J. Guertin, J. T. Lis, undergraduate thesis, Cornell University, Ithaca, NY (2012).
28. P. A. Emanuel, D. S. Gilmour, *Proc. Natl. Acad. Sci. U.S.A.* **90**, 8449 (1993).
29. B. Z. Ring, W. S. Yarnell, J. W. Roberts, *Cell* **86**, 485 (1996).

Acknowledgments: Supported by NIH grants GM25232 and HG004845 (J.T.L.) and a Howard Hughes Medical Institute fellowship (H.K.). Sequence data are in the Gene Expression Omnibus (GEO) database under accession number GSE42117. Part of this work is included in a broader U.S. patent application (12/554,472, "Genome-wide Method for Mapping of Engaged RNA Polymerases Quantitatively and at High Resolution"), which refers to variants of the method that we use here.

Supplementary Materials

www.sciencemag.org/cgi/content/full/339/6122/950/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S10
Tables S1 to S5
Source Codes for Analysis Scripts
References

27 August 2012; accepted 21 December 2012
10.1126/science.1229386

Little Emperors: Behavioral Impacts of China's One-Child Policy

L. Cameron,[1]* N. Erkal,[2] L. Gangadharan,[3] X. Meng[4]

We document that China's One-Child Policy (OCP), one of the most radical approaches to limiting population growth, has produced significantly less trusting, less trustworthy, more risk-averse, less competitive, more pessimistic, and less conscientious individuals. Our data were collected from economics experiments conducted with 421 individuals born just before and just after the OCP's introduction in 1979. Surveys to elicit personality traits were also used. We used the exogenous imposition of the OCP to identify the causal impact of being an only child, net of family background effects. The OCP thus has significant ramifications for Chinese society.

China's One-Child Policy (OCP) restricts the number of children that urban couples can have to one, with exceptions for those from ethnic minorities or with a severely disabled child. The policy has given rise to a land of "little emperors," whose parents dote on them exclusively. This has led to widespread concern within China about the social skills of this generation and the observation that these children tend to be more self-centered and less cooperative (*1*–*3*). This can be seen in developments such as employers including phrases like "no single children"

in job advertisements (*4*). In March 2007, 30 delegates at the Chinese People's Political Consultative Conference called on the government to abolish the policy (*5*). Their concerns centered on "social problems and personality disorders in young people."

The question about the behavioral consequences of not having siblings has been of interest in developmental psychology for many decades. China's OCP provides us with a natural experiment that allows us to identify the causal impact of being a single child. Prosocial development is believed to be shaped by parents as well as by social interactions with peers, including siblings (*6*). Different relationships with parents and lack of interaction with siblings have thus been identified in the psychology literature as two reasons why only children may develop differently from their counterparts with siblings (*7*). For example, parents of only children may be more responsive to their needs, which may produce a greater sense of security, confidence, and intellectual competence (*8*, *9*). Parents of only children may also be more able to interact with their children in ways that promote desirable development (*10*). More attention from parents can, however, come with downsides in terms of higher expectations and pressure to succeed in life (*11*). Only children are often viewed as disadvantaged as a result of "sibling deprivation," which may lead to their being self-centered, less cooperative, and less likely to get along with peers (*12*).

In this paper, we describe how we used techniques from experimental economics to measure behavioral differences between the pre- and post-OCP generations. Behavior in economic games has been widely shown to be correlated with actions outside the experimental setting (*13*–*18*). We investigated the impact of the OCP on altruism, trust, trustworthiness, attitudes toward risk, and competitiveness. The OCP can be thought of as a natural experiment that enables us to separate out the effect of being an only child from the effects of family background. In addition to our experimental results, personality survey questions revealed that the OCP cohorts are also substantially more pessimistic, less conscientious, and possibly more neurotic.

We conducted experiments with participants recruited from the general population of Beijing, where the policy is strictly enforced (*19*). The OCP was introduced in 1979. We required participants to be born in either 1975 or 1978 (our pre-OCP cohorts) or in 1980 or 1983 (our post-OCP cohorts). Participants were also required to have both parents with urban residency status (hukou) at the time of the participant's birth. This sampling ensured that those in the post-OCP cohorts were subject to the strictly enforced policy and that all the cohorts were similar with respect to their parental hukou status. The final sample consisted of 421 participants spread evenly across the birth cohorts 1975, 1978, 1980, and 1983, with gender balance within these cohorts.

A postexperiment participant survey collected demographic and socioeconomic information. We tested the representativeness of our sample by comparing it with the Beijing subsample of the National Bureau of Statistics 2009 Urban Household Survey data. Our sample was better educated than the general population but otherwise similar (table S1). We carefully controlled for education in the empirical specifications below.

Table 1 shows the proportion of only children in each birth cohort. This increased from 27% in the 1975 cohort to 91% for the cohort born in 1983. The average number of siblings decreased from 0.97 to 0.12 over the same period (*20*). The proportion of single children had increased before the 1979 introduction of the policy as a result of other noncompulsory population growth policies that were precursors to the OCP (fig. S1) (*21*). In this paper, we thus identify the behavioral impact of the OCP relative to the noncompulsory fertility policies implemented before the OCP.

Although in later years the policy undoubtedly reduced the size of extended families, extended family size differed only slightly between the pre- and post-OCP cohorts we studied. The number of cousins declined slightly from an average of 7.4 to 7.0 cousins (Table 1). We are thus identifying the impact of growing up as an only child, not the impact of having a smaller extended family.

The experiments we conducted are standard games from the economics literature. The dictator game (*22*) is designed to elicit the extent of altruism among participants. The trust game elicits the extent to which participants are able to trust one another and the extent of their trustworthiness (*23*). These games are explained in the supplementary materials.

In the risk game (*24*), each participant is given 100 yuan (which is approximately U.S. $15) and the choice to put any amount between 0 and 100 yuan into an "investment," which yields triple the amount invested with 50% probability and 0 with 50% probability. More risk-averse participants will invest less in the risky option. The outcome of the investment is decided by the flip of a coin at the completion of the session.

In the competition game, participants are asked to add up as many sets of five two-digit numbers as possible in 5 min (*25*). The numbers are randomly generated and presented in rows. Participants write the total in a blank box provided at the end of the row. Calculators cannot be used, but scrap paper is provided for handwritten calculations.

Participants are asked to choose between two different payment schemes. Option 1 is a piece rate which pays 5 yuan for every sum correctly completed. In option 2, payments are determined in a competitive way. Each participant is randomly and anonymously paired with someone else in the room. S/he is paid 10 yuan for every sum correctly completed if s/he completes more sums correctly than the person with whom s/he is paired, 5 yuan if both participants complete the same amount of correct sums, and 0 yuan if s/he loses the competition.

Figure 1 presents the unconditional differences in behavior between participants born before and after the OCP, for each of the games. The underlying data are presented in table S2 in the supplementary materials. Those born under the OCP shared slightly less of the endowment in the dictator game with the other player (40.1% of the endowment on average as compared to 43.4% sent by those born before the policy's introduction). The t test of difference in means is statistically insignificant ($P = 0.11$). OCP participants on average were less trusting, sending less to the other player (46.1 versus 50.6%) and returning less than those not born under the policy (30.4 versus 35.4%). Both of these differences are statistically significant.

OCP participants invested significantly less in the riskier investment (58.1 versus 66.4%). In the competition game, fewer OCP participants

Table 1. Demographic composition by cohort. We present means of variables by birth cohort and pre- and post-OCP. Column two reports means for those born in 1978 but excludes those born in 1978 who report they are an only child as a result of the policy.

	1975	1978[a]	1978	1980	1983	Pre-OCP	Post-OCP
% Only children	27.0	23.7	60.9	81.8	90.7	45.1	86.4
% With one sibling	56.0	62.7	32.2	17.2	5.6	43.3	11.2
% With two siblings	10.0	8.5	4.4	1.0	1.9	7.0	1.5
% With three sibling	4.0	3.4	1.7	0.0	0.9	2.8	0.5
% With four siblings	2.0	1.7	0.9	0.0	0.0	1.4	0
Average no. of siblings	0.97	0.97	0.50	0.19	0.12	0.72	0.16
% Being the first born	52.0	35.6	67.0	85.9	93.5	60.0	89.8
Average no. of cousins	6.93	7.95	7.78	7.14	6.90	7.4	7.0
Average no. of aunts and uncles	5.15	5.42	5.81	5.91	6.27	5.5	6.1
Age (years)	34.7	31.8	31.7	29.7	26.7	33.1	28.2
N	100	59	115	99	107	215	206

[1]Department of Econometrics, Monash University, Clayton, Victoria 3800, Australia. [2]Department of Economics, University of Melbourne, Victoria 3010, Australia. [3]Department of Economics, Monash University, Clayton, Victoria 3800, Australia. [4]Research School of Economics, College of Business and Economics, Australian National University, Australian Capital Territory 0200 Australia.

*To whom correspondence should be addressed. E-mail: lisa.cameron@monash.edu

chose to compete than those born before the policy (44.2 versus 51.8%). This difference is substantial but not statistically significant ($P = 0.12$).

Differences in competitiveness may reflect beliefs in one's ability. Participants were asked in which performance quintile they expected themselves to be relative to others in the room. There is no significant difference in this self-reported ranking between the pre- and post-OCP cohorts, although participants born under the OCP completed significantly more sums correctly than their counterparts (table S2). This is consistent with the findings in the literature that only children perform better academically (*26*).

The effects discussed above are simple mean differences. These could be due to differences in the demographic backgrounds of the pre- and post-OCP samples. To examine the OCP effects net of these observable characteristics, we estimated an econometric model in which we controlled for participants' gender and education, maternal education, and whether the individual was born in Beijing; in addition to the main variable of interest, whether the participant was born under the OCP. The summary statistics for the control variables are presented in table S3. Table 2 reports the estimation results. We find that even when the demographic and family background variables were controlled for, the unconditional effects we observed in Fig. 1 above persisted in terms of signs, magnitudes, and significance levels. The only difference from the unconditional estimates is that the regression results indicate that the lesser propensity for those who were born after the OCP to choose to compete becomes statistically significant once one controls for demographic characteristics.

The OCP indicator is defined according to birth cohort and hence is highly correlated with age. There is no evidence that age has a strong systematic effect on behavior in experiments of the type we conducted, particularly over the relatively small age range in our sample (*24*, *27*–*32*). Age was not a determinant of behavior within the pre- and post-OCP cohorts (table S8). Our results are also largely unaffected when we limit the potential for age effects by focusing on just the 1978 and 1980 cohorts (table S9). We thus conclude that age effects were not driving our results. Consistent with this result, marital status, being a parent, the age of parents, and people potentially becoming more capitalistic over time do not seem to explain our results either (tables S10 and S11).

Although the main effect of the OCP is to produce single children, the coefficients we estimate in Table 2 are not estimates of the effect of being a single child, because being a single child is not a unique characteristic of the after-OCP cohort. Many before-cohort individuals were also single children, and some children born after the policy were not single children (Table 1). The coefficient on the OCP indicator is thus the average behavioral effect of the OCP across the population.

If being a single child were exogenous, then the effect of growing up without siblings would be estimated consistently from

$$Y_i = \alpha + \beta X_i + \varphi Single_i + \upsilon_i \qquad (1)$$

where Y_i is the behavior of interest; $Single_i$ is an indicator for being an only child; α, β, and φ are coefficients to be estimated; and ν_i is a random error term. The coefficient of interest would be φ. Being a single child in the pre-policy cohort was, however, largely a choice of parents. Thus, the coefficient on $Single_i$ in Eq. 1 would pick up not only the effect of being a single child but also the effect of any omitted family background variables that influenced the probability of being an only child. This is a problem if the unobserved parental characteristics that make parents choose to have one child also affect individuals' behavioral outcomes via genetic or nurture channels, which is likely. The variable $Single_i$ is thus endogenous.

We can, however, exploit the exogeneity of the imposition of the OCP and use the indicator of whether one was born under the policy, D_i, to instrument for the endogenous variable $Single_i$. We can thus identify the causal effect of growing up without siblings as a result of the OCP. That is, we estimate the equations below using an instrumental variables (IV) approach. Equation 2b is the first-stage regression.

$$Y_i = \alpha + \beta X_i + \varphi Single_i + \upsilon_i \qquad (2a)$$

$$Single_i = \eta + \delta X_i + \kappa D_i + u_i \qquad (2b)$$

where η, δ, and κ are coefficients to be estimated; u_i is a random error term; and the other terms are defined as above. Provided that the instrument satisfies a few assumptions (see section 2.1 in the supplementary materials), the IV estimate of φ can be interpreted as the Local Average Treatment Effect (LATE) of growing up as only children because of the OCP (*33*).

The IV results are reported in Table 3. In every case where we identified a difference in behavior between the pre- and post-OCP cohorts, being an only child as a result of the OCP was also shown to have a significant causal impact on behavior. The coefficients on being an only child are over double the coefficients on the indicator of whether one was born under the OCP reported in Table 2. Individuals who were only children as a result of the policy sent on average 16 percentage points less of the endowment to the other

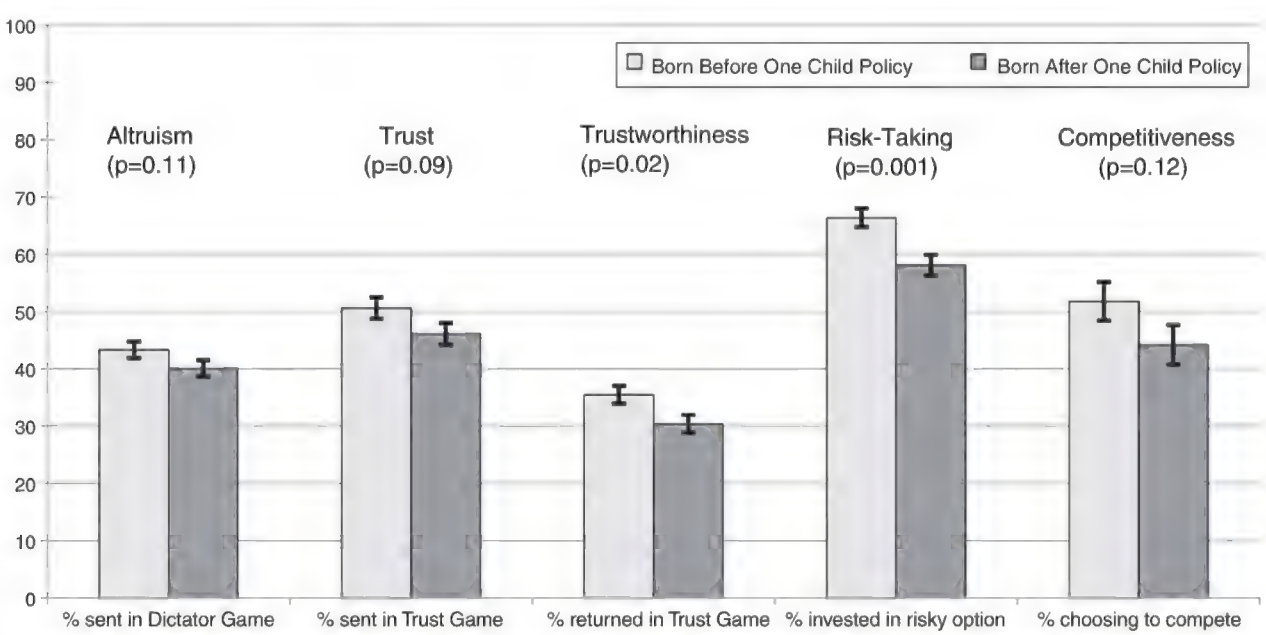


Fig. 1. Behavioral consequences of the OCP. Mean differences in behavior between participants born before and after the OCP are shown. The *P* values reported in parentheses are from *t* tests of differences in means. Error bars are mean ± SEM.

player and returned 11 percentage points less of what they received in the trust game, invested 20 percentage points less of the endowment in the risky option in the risk game, and were 20% less likely to choose the competitive option in the competition game. In terms of standard deviations (SD) of the dependent variable, being a single child is associated with a decrease of 0.58 SD of the percentage sent in the trust game, 0.44 SD of the percentage returned in the trust game, 0.75 SD of the percentage invested in the risky option, and 0.41 SD of the probability of competing. These are thus sizeable effects.

In the post-experiment survey, participants were asked a number of questions that sought to assess their personality type and outlook on life.

Table 2. Estimation results. We estimated $Y_i = \alpha + \beta X_i + \gamma D_i + \varepsilon_i$, where Y_i is a behavioral outcome observed in the experiments, and X_i is a vector of control variables, which includes participants' gender and education, maternal education, and whether the individual was born in Beijing. The main variable of interest is D_i, which equals 1 if an individual is born after the introduction of the OCP and 0 otherwise. The coefficient γ identifies any differences between those born before and after the policy and is our estimate of the behavioral impact of the OCP. Columns one to four report coefficients from Tobit estimation with lower censoring at 0 and upper censoring at 100. Column five presents marginal effects from a Probit estimation on whether to compete or not. Robust standard errors are shown in square brackets. *, **, *** indicate statistical significance at the 10, 5, and 1% levels, respectively.

	Altruism	Trust and trustworthiness		Risk-taking	Competitiveness
	Dictator game	Trust game		Risk game	Competition game
	% Sent	% Sent	% Returned	% Invested	Chose to compete
OCP	−2.834	−6.980**	−5.024*	−7.987**	−0.090*
	[2.275]	[3.116]	[2.728]	[3.213]	[0.050]
Male	−0.646	10.73***	0.694	3.086	0.072
	[2.299]	[3.120]	[2.687]	[3.152]	[0.049]
University or above	−14.37***	1.496	−15.34**	−14.030*	0.149
	[5.484]	[6.407]	[6.314]	[8.111]	[0.115]
3-year college	−13.28**	−3.989	−15.53**	−5.323	0.12
	[5.538]	[6.429]	[6.290]	[8.143]	[0.118]
Born in Beijing	−5.319*	−11.02**	4.788	5.391	0.082
	[2.977]	[4.656]	[4.485]	[4.096]	[0.075]
Mother with 3-year college	0.518	8.887*	4.594	8.123*	−0.009
	[3.784]	[4.669]	[4.253]	[4.220]	[0.067]
Mother with university or above	−4.363	6.983	−2.612	5.004	0.091
	[5.320]	[5.998]	[5.439]	[5.406]	[0.100]
Observations	410	408	398	416	419

Participants were asked, "What do you think are the chances that it will be sunny tomorrow? Please write a number from 0 to 100, where 0 means 'absolutely no chance' and 100 means 'absolutely certain.' " Responses to this question are widely used as an indicator of optimism (*34*). We found that those born under the OCP and those who grew up as single children as a result of the OCP were significantly less optimistic than others (table S15).

The post-experiment survey also implemented the Big Five Inventory, which contains 44 questions designed to categorize people in terms of openness (inventive/curious versus consistent/cautious), conscientiousness, extraversion, agreeableness, and neuroticism (sensitive/nervous versus secure/confident). These five broad factors together encompass most known personality traits, and the 44 questions are used to construct scores for each of these traits (*35*). We examined to what extent these traits were determined by the OCP and single-child status. Being born under the OCP and growing up as a single child as a result of the OCP are associated with a higher neuroticism score and a lower conscientiousness score (table S15).

These effects are not small: 0.69 SD of our optimism indicator, 0.52 SD of the conscientiousness score, and 0.71 SD of the neuroticism score. The result for neuroticism is consistent with the finding that positive sibling relationships moderate the relationship between stressful life events and internalizing behaviors (*36*). The finding that conscientiousness is lower is counter to the argument in the empirical literature in psychology that single children have a greater motivation to achieve, but it is consistent with Chinese parents' views of their only children (*37*).

Like the experimental results, our results for conscientiousness and optimism are qualitatively similar when estimated using just the 1978 and

Table 3. Causal impact of being an only child. We present results from IV estimation. We used an indicator of being born under the OCP as an instrument for being an only child. Columns one to four present the results of IV Tobit estimation. Column five presents marginal effects from IV Probit estimation. Robust standard errors are shown in square brackets. *, **, *** indicate statistical significance at the 10, 5, and 1% levels, respectively.

Dependent variables	Altruism (Dictator game; % sent)	Trust (Trust game; % sent)	Trustworthiness (Trust game; % returned)	Risk-taking (Risk game; % invested)	Competitiveness (Competition game; competing)
Only-child	−6.743	−15.998**	−11.34*	−19.002**	−0.203*
	[5.454]	[7.219]	[6.199]	[7.920]	[0.108]
Male	−0.829	10.604***	0.584	2.641	0.067
	[2.307]	[3.153]	[2.693]	[3.280]	[0.049]
University or above	−13.742**	3.064	−14.19**	−11.998	0.166
	[5.657]	[6.593]	[6.680]	[8.788]	[0.114]
3-year college	−12.573**	−2.323	−14.49**	−2.93	0.138
	[5.692]	[6.571]	[6.657]	[8.807]	[0.118]
Born in Beijing	−3.054	−5.914	8.442*	11.528**	0.142*
	[3.486]	[4.981]	[4.606]	[4.797]	[0.075]
Mother with 3-year college	1.238	10.396**	5.642	10.045**	0.011
	[3.837]	[4.721]	[4.251]	[4.605]	[0.069]
Mother with university or above	−3.418	8.522	−1.571	7.361	0.109
	[5.440]	[6.038]	[5.554]	[5.456]	[0.099]
Observations	410	408	398	416	419

1980 cohorts, and age is not a significant determinant within pre- and post-OCP periods. The results for neuroticism are less robust to checks for age effects (table S16).

Previous research has shown that noncognitive attributes such as conscientiousness, neuroticism, and optimism are important determinants of educational attainment, labor market outcomes, health, and marriage and divorce (*38*–*40*). Prosocial behavior is consistently seen to be an important determinant of social capital and plays a role in institutional development (*41*). A willingness to take risks is an important component of entrepreneurship (*17*). Our data show that being an only child as a result of the OCP is associated with taking less risk in the labor market (table S19).

Although our findings were obtained from a comparison of cohorts in Beijing born directly around the time of the policy's introduction, our results are generalizable to other urban areas of China where the OCP was strictly implemented. Previous work suggests that differences between only children and others in Beijing are similar to those in other urban areas (*26*). The effect of the policy on the behavior of people born long after the policy's introduction may, however, differ from what we found here, because later cohorts will have grown up with very limited extended family and in a society dominated by only children. Under such circumstances, we would expect that the policy's effect would, if anything, be magnified.

References and Notes

1. L. Lee, in *Child Care in Context: Cross Cultural Perspectives*, M. E. Lamb, K. Sternberg, Eds. (Lawrence Erlbaum, Hillsdale, NJ, 1992), pp. 355–392.
2. C. Fan, C. Wan, G. Lin, Q. Jin, *Psychol. Sci.* **17**, 70 (1994) (in Chinese).
3. Q. Wang, M. D. Leichtman, S. H. White, *Cognition* **69**, 73 (1998).
4. L. Chang, *Factory Girls: From Village to City in a Changing China* (Spiegel and Grau, New York, 2008).
5. *Consultative Conference: "The government must end the one-child rule."* AsiaNews.it, 16 March 2007, www.asianews.it/index.php?l=en&art=8757&size=A.
6. J. Piaget, *The Moral Judgment of the Child* (Free Press Paperbacks, New York, 1997).
7. L. Levy-Garboua, C. Meidinger, B. Rapoport, in *Handbook of the Economics of Giving, Altruism and Reciprocity*, S. C. Kolm, J. Mercier-Ythier, Eds. (Elsevier, Amsterdam, 2006), vol. I, pp. 545–613.
8. J. Bowlby, *Attachment and Loss* (Basic Books, New York, ed. 2, 1982), vol. 1.
9. J. Blake, *Demography* **18**, 421 (1981).
10. M. Lewis, M. C. Feiring, in *Families as Learning Environment for Children*, L. M. Laosa, I. E. Sigel, Eds. (Plenum, New York, 1982), pp. 115–145.
11. L. Roberts, P. Blanton, *J. Indiv. Psychol.* **57**, 125 (2001).
12. J. Blake, *Popul. Dev. Rev.* **7**, 43 (1981).
13. E. Fehr, A. Leibbrandt, *J. Public Econ.* **95**, 1144 (2011).
14. E. Liu, *Time to Change What to Sow: Risk Preferences and Technology Adoption Decisions of Cotton Farmers in China* (Princeton University Industrial Relations Section Working Paper No. 526, Princeton University, Princeton, NJ, 2008).
15. E. Liu, J. Huang, *Risk Preferences and Pesticide Use by Cotton Farmers in China* (University of Houston Working Paper, Houston, TX, 2012); www.class.uh.edu/Faculty/emliu/index.html.
16. D. Karlan, *Am. Econ. Rev.* **95**, 1688 (2005).
17. M. Benz, S. Meier, *Exp. Econ.* **11**, 268 (2008).
18. M. Castillo, R. Petrie, M. Torero, *Econ. Inq.* **48**, 266 (2010).
19. Materials and methods are available in the supplementary materials.
20. Some of the increase in the proportion of children without siblings in the 1978 cohort is due to parents not having had enough time to conceive a second child by the time of the introduction of the OCP in mid-1979. This issue is discussed in detail in the supplementary materials.
21. W. Lavely, R. Freedman, *Demography* **27**, 357 (1990).
22. R. Forsythe, J. Horowitz, N. Savin, M. Sefton, *Games Econ. Behav.* **6**, 347 (1994).
23. J. Berg, J. Dickhaut, K. McCabe, *Games Econ. Behav.* **10**, 122 (1995).
24. U. Gneezy, K. L. Leonard, J. A. List, *Econometrica* **77**, 1637 (2009).
25. M. Niederle, L. Vesterlund, *Q. J. Econ.* **122**, 1067 (2007).
26. T. Falbo, D. L. Poston Jr., *Child Dev.* **64**, 18 (1993).
27. L. Schechter, *J. Econ. Behav. Organ.* **62**, 272 (2007).
28. M. Sutter, M. Kocher, *Games Econ. Behav.* **59**, 364 (2007).
29. G. Charness, M. C. Villeval, *Am. Econ. Rev.* **99**, 956 (2009).
30. S. Andersen, G. Harrison, M. Lau, E. Rutstrom, *J. Econ. Behav. Organ.* **73**, 209 (2010).
31. G. Harrison, S. Humphrey, A. Verschoor, *Econ. J.* **120**, 80 (2010).
32. C. Engel, *Exp. Econ.* **14**, 583 (2011).
33. G. Imbens, J. Angrist, *Econometrica* **62**, 467 (1994).
34. See, for example, the U.S. Health and Retirement Survey.
35. For details of the questions used and how the index is calculated, see www.ocf.berkeley.edu/~johnlab/bfi.htm.
36. K. Gass, J. Jenkins, J. Dunn, *J. Child Psychol. Psychiatry* **48**, 167 (2007).
37. Y. Zhang, G. Geldolph, P. C. Cheung, S. Lau, *Social Science and Personality* **29**, 725 (2001).
38. J. Heckman, J. Stixrud, S. Urzua, *J. Labor Econ.* **24**, 411 (2006).
39. L. Borghans, A. Duckworth, J. Heckman, B. ter Weel, *J. Hum. Resour.* **43**, 972 (2008).
40. S. Lundberg, *Perspektiven Wirtschaftspolitik* **12**, 66 (2011).
41. E. Ostrom, J. Walker, *Trust and Reciprocity: Interdisciplinary Lessons for Experimental Research* (Russell Sage Foundation, New York, 2003).

Acknowledgments: We gratefully acknowledge funding from the Australian Research Council. We thank Simon Gaechter, Elaine Liu, and participants of various conferences and seminars for their useful comments on previous drafts. The data used in this paper will be made available from the ICPSR data depositary, number 33353. The Rumici data are available at rse.anu.edu.au/rumici. All authors contributed equally to this paper.

Supplementary Materials

www.sciencemag.org/cgi/content/full/science.1230221/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S4
Tables S1 to S19
Instructions to Participants
References

14 September 2012; accepted 18 December 2012
Published online 10 January 2013;
10.1126/science.1230221

Highly Recurrent *TERT* Promoter Mutations in Human Melanoma

Franklin W. Huang,[1,2,3]* Eran Hodis,[1,3,4]* Mary Jue Xu,[1,3,4] Gregory V. Kryukov,[1] Lynda Chin,[5,6] Levi A. Garraway[1,2,3]†

Systematic sequencing of human cancer genomes has identified many recurrent mutations in the protein-coding regions of genes but rarely in gene regulatory regions. Here, we describe two independent mutations within the core promoter of *telomerase reverse transcriptase* (*TERT*), the gene coding for the catalytic subunit of telomerase, which collectively occur in 50 of 70 (71%) melanomas examined. These mutations generate de novo consensus binding motifs for E-twenty-six (ETS) transcription factors, and in reporter assays, the mutations increased transcriptional activity from the *TERT* promoter by two- to fourfold. Examination of 150 cancer cell lines derived from diverse tumor types revealed the same mutations in 24 cases (16%), with preliminary evidence of elevated frequency in bladder and hepatocellular cancer cells. Thus, somatic mutations in regulatory regions of the genome may represent an important tumorigenic mechanism.

Systematic characterization of human cancer genomes has led to the discovery of a wide range of mutated genes that contribute to tumor development and progression. Most of the somatic mutations in tumors reside within the protein-coding regions of genes or at splice junctions. To determine whether tumor genomes harbor recurrent mutations outside of protein-coding regions, we systematically queried noncoding somatic mutations using published whole-genome sequencing data.

Analysis of whole-genome sequencing data from malignant melanomas (*1*, *2*) revealed two somatic *telomerase reverse transcriptase* (*TERT*) gene promoter mutations in 17 of 19 (89%) cases examined. The average sequence coverage at the *TERT* promoter locus was 30-fold in normal samples and 60-fold in tumor samples (fig. S1A).

[1]Broad Institute of Harvard and MIT, Cambridge, MA 02142, USA. [2]Department of Medical Oncology, Dana-Farber Cancer Institute, Boston, MA 02115, USA. [3]Harvard Medical School, Boston, MA 02115, USA. [4]Harvard–MIT Division of Health Sciences and Technology, Massachusetts Institute of Technology (MIT), Cambridge, MA 02139, USA. [5]Department of Genomic Medicine, M.D. Anderson Cancer Center, Houston, TX 77030, USA. [6]Institute for Applied Cancer Science, M.D. Anderson Cancer Center, Houston, TX 77030, USA.

*These authors contributed equally to this work.
†To whom correspondence should be addressed. E-mail: levi_garraway@dfci.harvard.edu

Each of these promoter mutations resulted in a cytidine-to-thymidine transition at a dipyrimidine motif indicative of ultraviolet (UV) light–induced damage (chr5, 1,295,228 C>T and 1,295,250 C>T; hereafter termed C228T and C250T, respectively), and both mutations localized within 100 base pairs (bp) of the *TERT* transcriptional start site (TSS) (mean allelic fraction, 0.32; range, 0.07 to 0.55) (table S1). We validated these mutations by means of polymerase chain reaction and Sanger sequencing tumor/normal sample pairs from both the discovery set (Fig. 1A and fig. S1, B and C) and an extension set of 51 additional melanoma tumor/normal sample pairs. Within this extension set, 33 tumors (65%) harbored one of the mutations. Moreover, the mutations were mutually exclusive in both the discovery and extension sets ($P = 5.4 \times 10^{-7}$, Fisher's one-sided exact test). Two tumors with a C228T transition also contained an adjacent C>T transition (at position chr5, 1,295,229), which is indicative of a dinucleotide CC>TT transition. Together, these *TERT* promoter mutations were observed in 50 of 70 (71%; 95% confidence interval: 59 to 82%, Clopper-Pearson method) melanomas examined (Fig. 1B and table S1).

Both C228T and C250T generated an identical 11-bp nucleotide stretch (5′-CCCCTTCCGGG-3′) containing a consensus binding site for E-twenty-six (ETS) transcription factors (GGAA, reverse complement) within the *TERT* promoter region. Because ETS transcription factors may become activated through dysregulation of mitogen-activated protein kinase (MAP kinase) signaling, we hypothesized that these promoter mutations might augment gene expression. To test this hypothesis, we used a reporter assay system in which the relevant portion of the mutant or wild-type *TERT* core promoter was cloned upstream of the firefly luciferase gene (*2*). Here, we tested both a core promoter fragment (–132 to +5 relative to the TSS) and the full core promoter (–200 to +73). In comparison to the wild-type *TERT* promoter, both mutations conferred approximately two- to fourfold increased transcriptional activity in five distinct cell line contexts (Fig. 1C and fig. S1D). Thus, each mutation was capable of augmenting transcriptional activity from the *TERT* promoter.

To investigate whether similar *TERT* promoter mutations occur in other cancer types, we examined sequencing data from this locus in 150 cell lines from the Cancer Cell Line Encyclopedia (CCLE) (*3*). Overall, 24 CCLE lines (16%) contained either C228T or C250T (mean allelic fraction, 0.61; range, 0.17 to 1.00) (table S1). An increased frequency in melanoma was again noted (five of six lines tested), with additional evidence suggesting possible heightened prevalence (>25%; one-sided 95% confidence interval) in bladder (three of three lines) and hepatocellular cancer cell lines (four of six lines) (Fig. 1D).

Several lines of evidence support the hypothesis that these promoter mutations may function as driver events that contribute to oncogenesis through TERT dysregulation and undergo positive selection, at least in human melanoma. First, the *TERT* promoter mutations showed a combined frequency that exceeded those of *BRAF* and *NRAS* mutations, which activate known melanoma driver oncogenes (*4*, *5*). In an analysis restricted to somatic mutations present at an allelic fraction of 0.2 or greater [to reduce artifacts of mutation calling (*1*)], the four most recurrent melanoma nucleotide substitutions included *BRAF* [chr7, 140,453,136 A>T (V600E)], *NRAS* [chr1, 115,256,529 T>C (Q61R)], and the *TERT* core promoter mutations C228T and C250T. Second, although highly recurrent, C228T and C250T occurred in a wholly mutually exclusive fashion. This suggests the possibility that the mutations might be functionally redundant. Third, the absence of other recurrent somatic mutations in the 3 kb upstream of the *TERT* transcription start site in the queried melanomas (*1*) coupled with the absence of the described *TERT* promoter mutations in 24 lung adenocarcinomas with comparably high somatic mutation rates (*6*) reduces the possibility that these recurrent *TERT* promoter mutations are solely due to an increased background mutation rate at this locus.

Although the role of telomerase in tumorigenesis is well established, details regarding its dysregulation in cancer cells remain incompletely understood, particularly in melanoma (*7*). The *TERT* promoter mutations identified here may link telomerase gene regulation and tumorigenic activation in this malignancy. The high prevalence of C228T and C250T suggests that these *TERT* promoter mutations may comprise early genetic events in the genesis of melanoma and other cancer types. Although TERT expression alone is not sufficient to bypass oncogene-induced senescence, genomic *TERT* activation may potentiate mechanisms by which melanocytes achieve immortalization in the setting of oncogenic mutations (*8*). These results therefore suggest that renewed efforts to develop clinically effective telomerase inhibitors may be warranted.

At the same time, promoter mutations likely represent only one potential mechanism of *TERT* reactivation in a subset of human cancers. Indeed, recurrent chromosomal copy gains spanning the *TERT* locus have been described previously for several cancers, including melanoma (*9*, *10*).

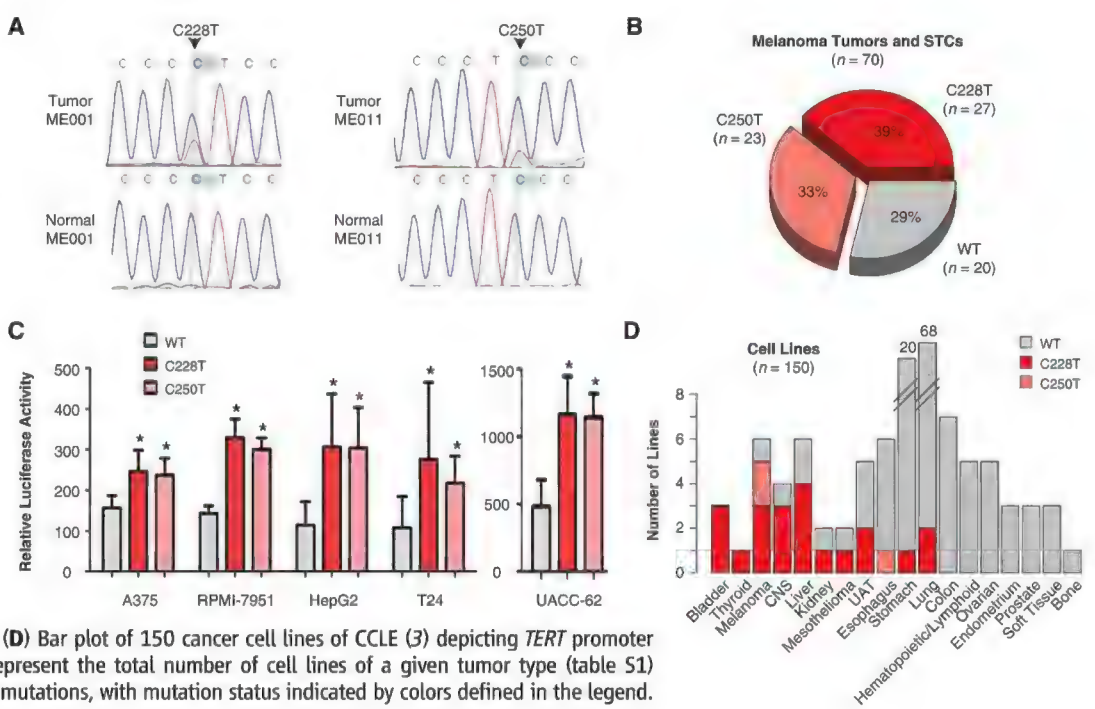


Fig. 1. Identification of *TERT* promoter mutations in melanoma and cancer cell lines. **(A)** Sequence chromatograms of matched tumor and normal DNA representing somatic mutations chr 5 [1,295,228 C>T (C228T)] and chr 5 [1,295,250 C>T (C250T)] in the *TERT* promoter locus. **(B)** Pie chart of C228T and C250T somatic mutation status in 70 surveyed melanoma tumors and short-term cultures. Sum of percentages is greater than 100% because of rounding. **(C)** Luciferase reporter assays for transcriptional activity from the *TERT* core promoter (–200 to +73) with either the C228T or C250T mutation compared with wild-type promoter in A375, RPMI-7951, UACC-62, T24, or HepG2 cell lines. The results depicted are the average of at least three independent experiments. Values are mean ± SD; $^*P < 0.05$. **(D)** Bar plot of 150 cancer cell lines of CCLE (*3*) depicting *TERT* promoter mutation status. Individual bars represent the total number of cell lines of a given tumor type (table S1) interrogated for C228T and C250T mutations, with mutation status indicated by colors defined in the legend.

Highly recurrent somatic mutations within a cancer gene promoter region have not previously been described. Similarly, the de novo mutational generation of transcription factor binding motifs in tumor genomes was heretofore unknown, although an ETS transcription factor binding motif was previously associated with a single-nucleotide polymorphism insertion at the MMP-1 locus (*11*). Together, these findings raise the possibility that recurrent somatic mutations involving regulatory regions, in addition to coding sequences, may represent important driver events in cancer.

References and Notes

1. M. F. Berger *et al.*, *Nature* **485**, 502 (2012).
2. Materials and methods are available as supplementary materials on *Science* Online.
3. J. Barretina *et al.*, *Nature* **483**, 603 (2012).
4. E. Hodis *et al.*, *Cell* **150**, 251 (2012).
5. M. Krauthammer *et al.*, *Nat. Genet.* **44**, 1006 (2012).
6. M. Imielinski *et al.*, *Cell* **150**, 1107 (2012).
7. D. C. Bennett, *Melanoma Res.* **21**, 27 (2008).
8. C. Michaloglou *et al.*, *Nature* **436**, 720 (2005).
9. A. Zhang *et al.*, *Cancer Res.* **22**, 6320 (2000).
10. C. Pirker *et al.*, *Melanoma Res.* **13**, 483 (2003).
11. J. L. Rutter *et al.*, *Cancer Res.* **58**, 5321 (1998).

Acknowledgments: We thank S. N. Wagner, D. Schadendorf, J. A. Wargo, and D. S. B. Hoon for contribution of samples; K. Cibulskis for assistance with whole-genome sequence data analysis; and E. Nickerson for coordination of validation/extension studies. This work was supported by NIH grant T32 CA009172 (F.W.H.), the Mittelman Family Fellowship (F.W.H.), National Institute of General Medical Sciences grant T32GM07753 (E.H.), the American Cancer Society (M.J.X.), NIH New Innovator Award DP2OD002750 (L.A.G.), National Cancer Institute grant R33CA126674 (L.A.G.), the Novartis Institutes for Biomedical Research (G.V.K. and L.A.G.), the Melanoma Research Alliance (L.A.G.), and the Starr Cancer Consortium (L.C. and L.A.G.). L.A.G. is on the Scientific Advisory Board of Millennium Pharmaceuticals, is a paid consultant for and holds equity in Foundation Medicine, and is a paid consultant for and sponsored research recipient from Novartis. The BAM files for the CCLE Whole Genome Sequencing data corresponding to the TERT locus can be accessed at ftp://tertguest:broad@ftp.broadinstitute.org.

Supplementary Materials

www.sciencemag.org/cgi/content/full/science.1229259/DC1
Materials and Methods
Fig. S1
Table S1
References (*12*, *13*)

24 August 2012; accepted 2 January 2013
Published online 24 January 2013;
10.1126/science.1229259

TERT Promoter Mutations in Familial and Sporadic Melanoma

Susanne Horn,[1,2] Adina Figl,[1,2] P. Sivaramakrishna Rachakonda,[1] Christine Fischer,[3] Antje Sucker,[2] Andreas Gast,[1,2] Stephanie Kadel,[1,2] Iris Moll,[2] Eduardo Nagore,[4] Kari Hemminki,[1,5] Dirk Schadendorf,[2]*† Rajiv Kumar[1]*†

Cutaneous melanoma occurs in both familial and sporadic forms. We investigated a melanoma-prone family through linkage analysis and high-throughput sequencing and identified a disease-segregating germline mutation in the promoter of the *telomerase reverse transcriptase* (*TERT*) gene, which encodes the catalytic subunit of telomerase. The mutation creates a new binding motif for Ets transcription factors and ternary complex factors (TCFs) near the transcription start and, in reporter gene assays, caused up to twofold increase in transcription. We then screened the *TERT* promoter in sporadic melanoma and observed recurrent ultraviolet signature somatic mutations in 125 of 168 (74%) of human cell lines derived from metastatic melanomas, 45 of 53 corresponding metastatic tumor tissues (85%), and 25 of 77 (33%) primary melanomas. The majority of those mutations occurred at two positions in the *TERT* promoter and also generated binding motifs for Ets/TCF transcription factors.

The identification of germline mutations that cosegregate with disease in cancer-prone families often provides genetic and mechanistic insights into the more common, sporadically arising cancers. In a study of cutaneous melanoma, the most malignant skin cancer, we investigated a large pedigree with 14 related melanoma patients who were not carriers of germline mutations in *CDKN2A* or *CDK4*, two known melanoma genes (Fig. 1). Multipoint linkage analysis showed a possible 2.2-Mb linkage region on chromosome 5p with maximal logarithm of the odds ratio for linkage scores of 2.35 at rs1379917 and 2.45 at rs1968011. Target-enriched high-throughput sequencing (HTS) of the region was carried out on constitutional DNA from the four affected and four unaffected members of the family with an average coverage between 55- and 108-fold (table S1) (*1*). The HTS data revealed a single promoter variant, three intronic variants, and three nongene variants previously unknown and unique to the DNA sequences of the affected individuals (table S2). The disease segregating variants, seven in total, were validated by Sanger sequencing of DNA from the individuals sequenced by HTS and of DNA from additional unaffected members of the family. The new variants were also detected in an unaffected member (754, table S3), who was 36 years old and carried multiple nevi. DNA from affected individuals other than those sequenced by HTS was not available for testing.

Of the seven unique variants identified, one variant (T>G), was located in the promoter at –57 base pairs (bp) from ATG translation start site of the *telomerase reverse transcriptase* (*TERT*) gene. The *TERT* gene encodes the catalytic reverse transcriptase subunit of telomerase, the ribonucleoprotein complex that maintains telomere length. The nucleotide change in the sequence CCTGAA>CCGGAA creates a new binding motif for Ets transcription factors, with a general recognition motif GGA(A/T). Beyond the general motif for Ets transcription factors, the familial mutation also generates a binding motif, CCGGAA, for the ternary complex factors (TCFs) Elk1 and Elk4 (*2*, *3*). To exclude the possibility that the detected promoter mutation in *TERT* is a common germline variant, we screened germline DNA from 140 sporadic melanoma cases and 165 healthy controls, and none carried the variant. Screening of DNA from index cases from 34 Spanish melanoma families also did not show any mutations. No carriers were found in dbSNP and the 1000 Genomes databases (data available for 18 individuals were obtained from Ensembl).

The familial mutation in the *TERT* promoter was in complete allelic linkage with a common polymorphism rs2853669 (G>A) at –246 bp upstream from the ATG start site (table S3). In previous work, this polymorphism was reported to disrupt an Ets binding site, and it was associated with low telomerase activity in patients with non–small cell lung cancer (*4*). In luciferase reporter gene assays, we found that the activity of constructs containing the mutation at –57 bp of the *TERT* promoter was increased 1.5-fold and 1.2-fold over the wild-type construct in Ma-Mel-86a and human embryonic kidney (HEK) 293T cells, respectively. A construct with both the *TERT* mutation and the variant allele of the rs2853669 polymorphism showed a 2.2-fold increase in promoter activity in Ma-Mel-86a and and 1.3-fold increase in HEK293 cells (mean from three measurements; details in supplementary text and fig. S1).

The germline occurrence of the promoter mutation, creating an Ets/TCF motif, can result in modification of *TERT* expression in all tissues expressing Ets/TCF. Highest staining for the TCF Elk1 protein has been reported in female-specific tissues, such as ovary and placenta. The increased expression of TCF Elk1 protein in female-specific tissues may cause gender-related differences in

[1]Division of Molecular Genetic Epidemiology, German Cancer Research Center, Im Neuenheimer Feld 580, 69120 Heidelberg, Germany. [2]Department of Dermatology, University Hospital Essen, 45122 Essen, Germany. [3]Institute of Human Genetics, University of Heidelberg, 69120 Heidelberg, Germany. [4]Department of Dermatology, Instituto Valenciano de Oncologia, Valencia, Spain. [5]Center for Primary Health Care Research, Lund University, Malmö, Sweden.

*To whom correspondence should be addressed. E-mail: r.kumar@dkfz.de (R.K.); dirk.schadendorf@uk-essen.de (D.S.)
†These authors contributed equally to this work.

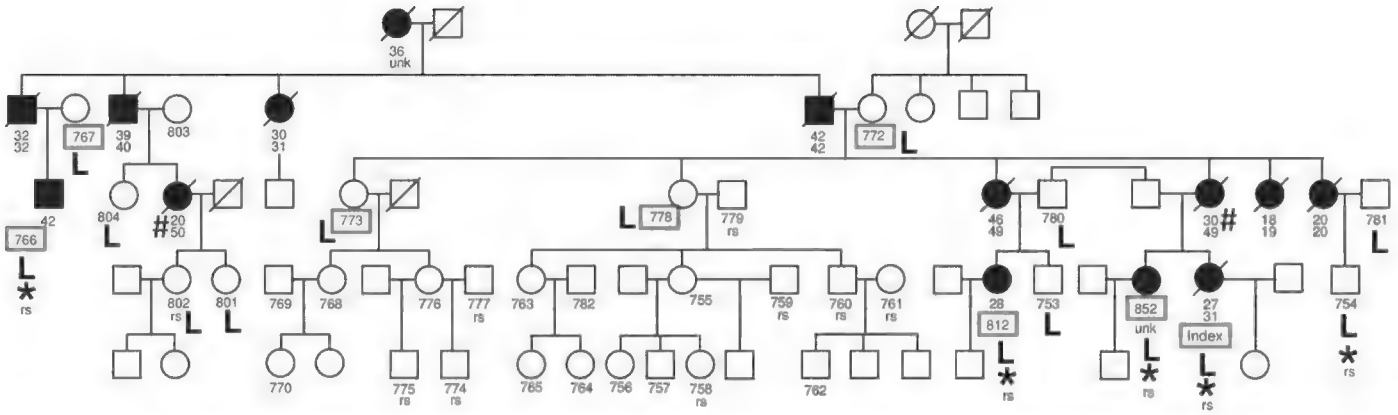


Fig. 1. Pedigree of melanoma-prone family. Four generations were affected by melanoma (solid symbols; circles represent females, and squares represent males). After linkage analysis carried out on 15 family members (L), HTS was performed on four affected and four unaffected individuals (boxed samples). A mutation in the *TERT* promoter was identified in all affected members and one unaffected individual (stars). Strikethrough symbols indicate deceased individuals. Two-digit numbers are age at onset of melanoma and age at death; Unk, unknown; Rs, rs2853669 observed in heterozygous form; three-digit numbers, DNA available; #, affected by other cancers; and index, index patient.

cancer susceptibility among carriers of the *TERT* mutation (*5*) (supplementary text). Two affected members of the family developed several different types of cancer (marked with # in Fig. 1). One affected individual presented with ovarian cancer at age 27 and melanoma at age 30. Another individual was diagnosed with melanoma at age 20; later she developed ovarian cancer, renal cell carcinoma, bladder cancer, mammary carcinoma, and finally bronchial carcinoma, leading to her death at age 50.

The mutation in the melanoma-prone family prompted us to screen melanoma cell lines derived from sporadic cases of metastatic melanoma. None of the cell lines carried the mutation detected in the family. However, we identified recurrent ultraviolet (UV)–signature mutations in the *TERT* core promoter in 74% (125 of 168) of the cell lines. The mutations were located within a 49-bp region starting from –100 bp upstream of the ATG start site (Table 1, Fig. 2, fig. S2, and table S4). There were two frequent mutations at –124 bp (G>A; C>T on opposite strand) and –146 bp (G>A); these mutations were mutually exclusive and occurred in 27 and 38% of cell lines, respectively. Two tandem GG>AA (CC>TT) mutations at positions –124/–125 bp and –138/–139 bp were observed at a frequency of 9%. The tandem mutation at positions –138/–139 bp could also be generated by a single-base mutation at –138 bp, because the base change at –139 bp has been reported as a rare polymorphism (rs35550267). The two most frequent single-base mutations as well as the two tandem mutations also result in the creation of Ets/TCF binding motifs.

Mutations were confirmed in 45 of 53 (85%) available metastasized tumors corresponding to the cell lines. The somatic nature of the mutations was shown by the absence of mutations in corresponding DNA from peripheral blood mononuclear cells available from 23 patients. Somatic mutations in the *TERT* promoter were more frequent than the *BRAF* mutations (53%, 90 of 169), *CDKN2A* alterations (50%, 84 of 169), and *NRAS* mutations (23%, 38 of 169; fig. S3). The occurrence of concomitant mutations in the *TERT* promoter and *BRAF* was more frequent (47%) than by random chance (40%) with an odds ratio (OR) of 3.2 [95% confidence interval (CI) 1.3 to 8.2]. Concomitant mutations in *TERT*, *BRAF*, and *CDKN2A* were observed in 30% of cell lines compared with the expected frequency of such occurrence of 9% (OR 5.6, 95% CI 2.4 to 13.8). The high recurrence and specificity of the *TERT* promoter mutations, together with the preliminary evidence from reporter assays that they have a functional effect on transcription, suggest that these mutations are driver rather than passenger events. Extensive functional studies will be required to validate this hypothesis.

Table 1. Most-frequent *TERT* core promoter mutations in screened metastatic melanoma cell lines and paraffin embedded primary tumors. A total of 169 cell lines were screened. Amplification for the *TERT* promoter failed for one cell line. Of 168 cell lines examined, 125 carried recurrent mutations. Of 77 primary melanomas examined, 24 carried recurrent mutations, and one carried a rare mutation (table S5). Seven rare mutations occurred at other sites in less than 2% of samples. Details of all mutations and polymorphisms are given in tables S4 and S5. Matched normal control DNA corresponding to 23 cell lines did not show mutations. For primary tumors, matched normal control DNA was not available.

Position (hg19) and variant	Distance to start (bp)	Cell lines	%	Primaries	%
1,295,228 G>A	–124	46	27.4	7	9.1
1,295,228 and 1,295,229 GG>AA	–124, –125	7	4.2	4	5.2
1,295,242 and 1,295,243 (rs35550267) GG>AA	–138, –139	8	4.8	8	10.4
1,295,250 G>A	–146	64	38.1	5	6.5

The *TERT* promoter mutations were also detected in 25 out of 77 (33%) paraffin embedded primary melanoma tumors (Table 1 and table S5) at –124 bp (7/77; 9%) and –146 bp (5/77; 7%). Four primary tumors carried the GG>AA tandem mutations at –124/–125 bp, and eight primary tumors carried the GG>AA tandem mutations at –138/–139 bp. Reduced sensitivity to detect mutations in paraffin-embedded primary tumors because of contaminating normal cells cannot be ruled out. Primary tumors harbored five additional mutations in the *TERT* promoter, which were not present in metastases, and those did not generate Ets/TCF binding motifs. We also screened DNA extracted from 25 melanocytic nevi and only one carried a mutation at –101 bp, which did not create an Ets/TCF motif. For both primary tumors and melanocytic nevi, matched normal control DNA was not available for testing.

The *TERT* coding region has been reported to be somatically mutated in 1% of cancers (14 cancer types, 1271 unique samples) (*6*). Mutations creating Ets/TCF binding motifs in the *TERT* promoter in melanoma have not been described in earlier sequencing projects.

TCFs are a subfamily of Ets transcription factors; two members of this subfamily, Elk1 and Elk4, are downstream targets of *BRAF* and regulate the expression of many genes (*7*–*11*). Conceivably, TCF may represent a link between telomerase activity and the frequent *BRAF* activating mutations in melanoma (fig. S4) (*12*, *13*). Lastly, whether *TERT* promoter mutations occur in other cancer types remains to be determined. We did not detect these mutations in a screen of 22 esophageal squamous cell carcinomas, but further analyses are warranted.

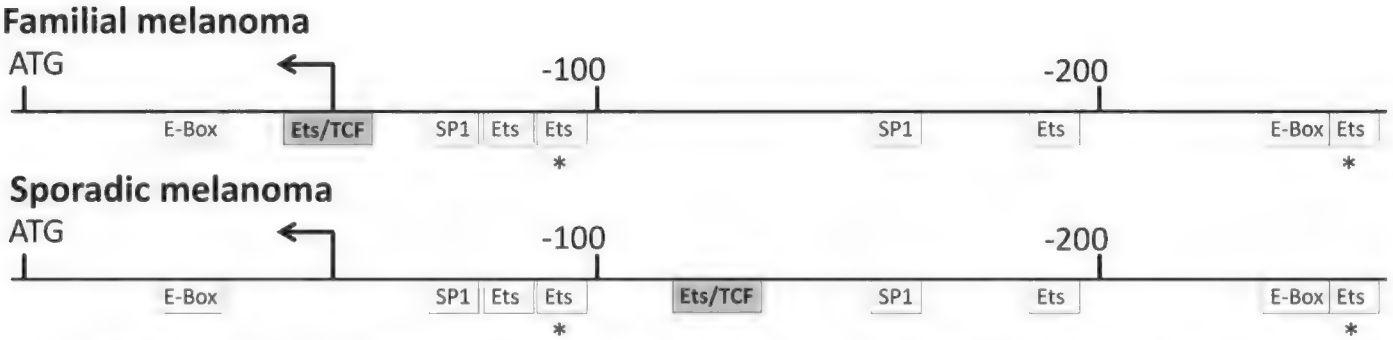


Fig. 2. The *TERT* core promoter in melanoma. Mutations creating Ets/TCF binding motifs were found in affected family members (–57 bp) immediately next to the transcription start site and in sporadic metastatic melanoma (–124 to –149 bp; sequence details in fig. S2). Binding sites for c-Myc (E-Box), SP1, and Ets transcription factors are known to exist in the wild-type *TERT* promoter. Ets2 binding was reported for Ets2 sites at –99 and –243 bp (stars) (*4*). The plus strand of DNA is shown.

References and Notes

1. Materials and methods are available as supplementary materials on *Science* Online.
2. Y. Akiyama, TFSearch, www.cbrc.jp/htbin/nph-tfsearch (1995).
3. P. Shore, A. D. Sharrocks, *Nucleic Acids Res.* **23**, 4698 (1995).
4. D. Xu, J. Dwyer, H. Li, W. Duan, J.-P. Liu, *J. Biol. Chem.* **283**, 23567 (2008).
5. M. Uhlen *et al.*, *Nat. Biotechnol.* **28**, 1248 (2010).
6. S. A. Forbes *et al.*, *Nucleic Acids Res.* **39**, D945 (2011).
7. V. N. Rao, E. S. Reddy, *Oncogene* **9**, 1855 (1994).
8. A. J. Whitmarsh, P. Shore, A. D. Sharrocks, R. J. Davis, *Science* **269**, 403 (1995).
9. W. Lee, E. B. Keller, *J. Mol. Biol.* **220**, 599 (1991).
10. G. C. Kujoth, D. F. Robinson, W. E. Fahl, *Cell Growth Differ.* **9**, 523 (1998).
11. R. Janknecht, W. H. Ernst, A. Nordheim, *Oncogene* **10**, 1209 (1995).
12. H. Davies *et al.*, *Nature* **417**, 949 (2002).
13. R. Kumar *et al.*, *Clin. Cancer Res.* **9**, 3362 (2003).

Acknowledgments: We thank Eurofins MWG Operon (Ebersberg, Germany) for sequencing service and M. Kircher and U. Stenzel for comments on sequence analysis. The study was partially supported by a grant from Deutsche Forschungsgemeinschaft (DFG). The mutation data reported in this manuscript are deposited in www.sanger.ac.uk/genetics/CGP/cosmic/.

Supplementary Materials

www.sciencemag.org/cgi/content/full/science.1230062/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S7
Tables S1 to S7
References (*14*–*33*)

11 September 2012; accepted 11 January 2013
Published online 24 January 2013;
10.1126/science.1230062

Increases in Adult Life Expectancy in Rural South Africa: Valuing the Scale-Up of HIV Treatment

Jacob Bor,[1,2]* Abraham J. Herbst,[1] Marie-Louise Newell,[1,3] Till Bärnighausen[1,2]

The scale-up of antiretroviral therapy (ART) is expected to raise adult life expectancy in populations with high HIV prevalence. Using data from a population cohort of over 101,000 individuals in rural KwaZulu-Natal, South Africa, we measured changes in adult life expectancy for 2000–2011. In 2003, the year before ART became available in the public-sector health system, adult life expectancy was 49.2 years; by 2011, adult life expectancy had increased to 60.5 years—an 11.3-year gain. Based on standard monetary valuation of life, the survival benefits of ART far outweigh the costs of providing treatment in this community. These gains in adult life expectancy signify the social value of ART and have implications for the investment decisions of individuals, governments, and donors.

For most of the 20th century, life expectancy increased in nearly every part of the world (*1*). However, from the late 1980s, the HIV epidemic led to a reversal of this trend in southern Africa, with a large rise in mortality among working-age adults (*1*–*3*). In South Africa, life expectancy at age 15 declined from 67.4 years in 1990 to 58.7 years in 2009; and in Swaziland, from 68.1 to 53.4 years (*2*). In addition to the direct loss of life, these declines in adult life expectancy had profound negative effects on households, communities, and governments, including declines in household wealth; large increases in the number of orphans; the loss of skilled workers, including teachers, doctors, and government officials; and the interruption of intergenerational transmission of knowledge and norms (*4*).

In the early 2000s, southern African nations began to disburse mass antiretroviral therapy (ART) for HIV through public-sector treatment programs, often with support from international donors. Using a combination of three or more drugs, ART interrupts HIV replication, enables immune recovery, and improves survival among people with HIV (*5*). Population-level declines in HIV-related and all-cause mortality have been documented in South Africa (*6*–*9*), Malawi (*10*), and other countries receiving financial assistance for HIV programs from the U.S. government (*11*). However, the impact of ART on population-level adult life expectancy in highly affected communities has not been quantified.

Life expectancy summarizes age patterns of mortality in a single statistic and is commonly used to compare differences in mortality across populations and over time (*1*–*3*). Because HIV predominantly affects working-age adults, adult life expectancy is of particular interest for governments and donors, as well as for individuals and households, whose plans for the future will be influenced by changes in the anticipated length of life. Adult life expectancy is defined as the mean age to which a 15-year-old could expect to live if subjected to the full pattern of age-specific mortality rates observed for a population over a particular period of time. Because future mortality rates are unknown, adult life expectancy cannot be interpreted as the average age to which a cohort will live, except in the limited case in which age-specific mortality rates remain constant into the future. Adult life expectancy is best interpreted as a summary indicator of the mortality experience in a population at a given time.

This paper documents the impact of South Africa's public-sector scale-up of ART on adult life expectancy in a large population cohort in

[1]Africa Centre for Health and Population Studies, University of KwaZulu-Natal, Post Office Box 198, Mtubatuba, KwaZulu-Natal 3935, South Africa. [2]Harvard School of Public Health, Harvard University, 677 Huntington Avenue, Boston, MA 02115, USA. [3]Institute of Child Health, University College London, 30 Guilford Street, London WC1N 1EH, UK.

*To whom correspondence should be addressed. E-mail: jbor@hsph.harvard.edu

rural KwaZulu-Natal, a setting with very high HIV prevalence (*12*). The life expectancy of HIV patients who begin ART before they fall severely ill and who subsequently adhere to their ART regimens approaches the life expectancy of people who are HIV-negative (*13*). However, extrapolation to population-level life expectancy is not straightforward, owing to the difficulties of measuring treatment coverage (*14*), adherence and retention (*15*), and survival for patients presenting at later stages of HIV disease (*16*). In addition, ART programs may have spillover effects on other aspects of health systems, which may in turn affect non-HIV mortality, although the direction and magnitude of such effects have yet to be established (*17, 18*). ART scale-up may also have other spillover effects on mortality, through changes in rates of depression and suicide, health care–seeking for other conditions, HIV risk compensation, and risk behaviors linked to survival expectations such as substance abuse and violence. By focusing on population-level life expectancy, we automatically account for any spillover effects on contemporaneous non-HIV mortality.

Existing estimates of population-level life expectancy in the era of ART are based on demographic models (*1, 2, 19–21*), which rely on a range of assumptions. For example, the South African government uses United Nations East Asia model life tables to infer age-specific mortality rates (*20*). In contrast to modeling approaches, we directly measured dates of death using individual-level data from a large community-based population surveillance system.

The study population included all adult resident and nonresident members of all households in a 434-km^2 surveillance area in rural KwaZulu-Natal. Data on births and deaths were collected from 2000 through 2011 via semiannual household survey visits, with response rates >99% (*22, 23*). This Health and Demographic Surveillance System is maintained by the Africa Centre for Health and Population Studies (Africa Centre), a research center funded by the Wellcome Trust and affiliated with the University of KwaZulu-Natal. The Africa Centre's population surveillance is a dynamic (open) cohort (*24*), in which individuals are observed from the date when they join a household in the surveillance area. To account for complex patterns of cyclical migration, individuals are observed regardless of whether they reside in the surveillance area, provided that they are members of a household under surveillance. Individuals were included in this analysis so long as their death would have been observed had it occurred. The study included a total of 101,286 persons, of whom about 60,000 were aged 15 years or older and under surveillance in any given year.

The community is largely rural and is located in one of the poorest districts in South Africa (*25*). HIV prevalence is very high; 29% of adults are HIV-positive, with about half of women aged 30 to 49 and about one-third of men aged 35 to 49 living with HIV (*12*). In the early 2000s, over half of all deaths in this community were attributable to HIV (*7*).

In 2004, South Africa began to provide ART for HIV-infected adults at government clinics and hospitals, with the goal of achieving universal coverage for all individuals meeting disease-stage eligibility criteria. The public-sector HIV treatment program serving the Africa Centre surveil-

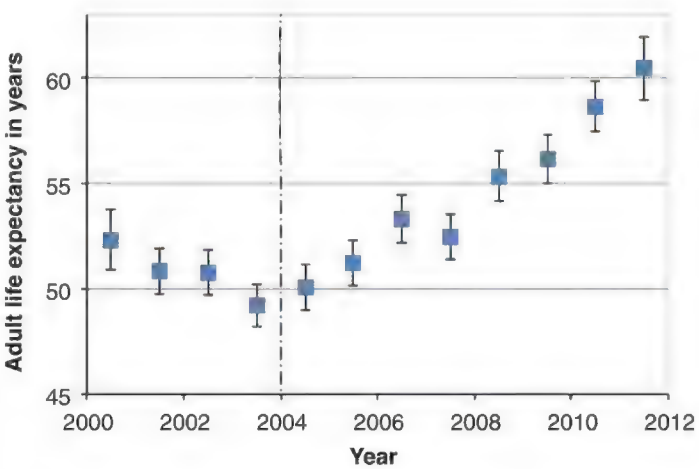


Fig. 1. Adult life expectancy, 2000–2011. Adult life expectancy is the mean age to which a 15-year-old could expect to live if subjected to the full pattern of age-specific mortality rates observed in a population for a given period of time. Annual estimates of adult life expectancy (blue squares) are shown for each year, 2000 to 2011, with 95% CIs. Public-sector provision of ART to adults in this community began in 2004, as indicated by the vertical line.

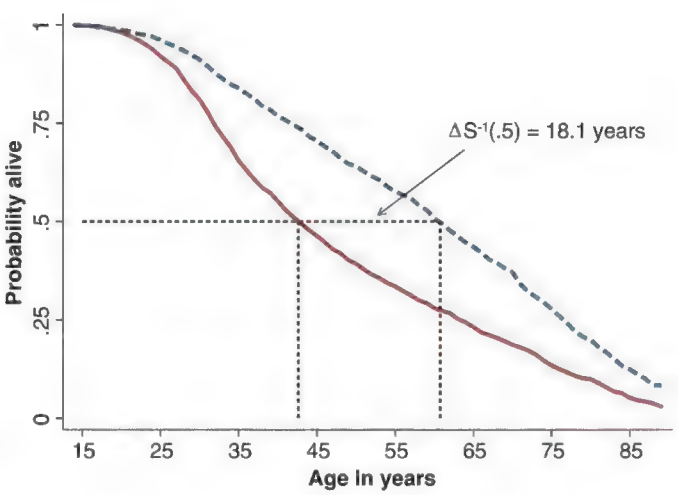


Fig. 2. Survival curves for 2003 and 2011. Kaplan-Meier survival curves for 2003 (solid red line) and 2011 (broken blue line) were estimated for the population under surveillance. Each curve displays the probability that someone would be alive at a given age if subjected to the full pattern of age-specific mortality rates observed in that year. Conditional on survival to 15 years, the median length of life was 42.6 years in 2003 and 60.7 years in 2011.

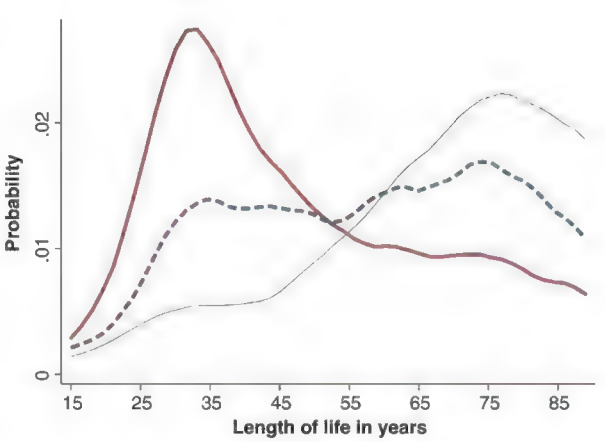


Fig. 3. Distributions of lengths of life are presented for 2003 (solid red line) and 2011 (broken blue line). The thin green line displays the distribution of HIV-cause–deleted lengths of life for 2001–2010, which are based on mortality rates that exclude HIV-related deaths. The proportion of deaths occurring in young adulthood declined between 2003 and 2011, but there was still evidence of excess HIV-related mortality among young adults in 2011 (by comparison to the HIV-cause–deleted distribution of lengths of life).

lance area first enrolled patients in September 2004. The program is administered by the Department of Health, is led by nurses at community-based clinics, and is largely publicly financed. The Africa Centre has supported this program since its inception with funding from the U.S. President's Emergency Plan for AIDS Relief (PEPFAR); in particular, the Africa Centre has provided health worker training, salary support, and technical assistance, such as the management of an electronic database.

By 1 July 2011, 12.6% of adults aged 15 and older residing in this community had sought care in the public-sector HIV treatment program; 7.0% had initiated ART, representing more than a third of all HIV-infected adults living in the community (*26*) (fig. S1 and table S1). Because of the high cost of antiretroviral drugs and low levels of health insurance coverage, private-sector ART utilization has always been very low in this community. The public-sector scale-up of ART was a clear shift in the therapeutic options available to people with HIV (*27*).

Trends in adult life expectancy both before (2000–2003) and after (2004–2011) ART became available in the public sector were analyzed using a continuous-time approach (*28, 29*). Survival curves were estimated separately for each year, using the Kaplan-Meier estimator (*30*). Adult life expectancy, or mean length of life, was calculated as the area under the estimated survival curve. For single-year estimates of adult life expectancy, data at ages >95 years were sparse. To avoid bias due to variation in the age of the oldest cohort member across calendar years, we estimated life expectancy between ages 15 and 95 years. Ninety-five percent confidence intervals (CIs) were calculated for each annual estimate (*28*). In the pooled 2000–2011 data, the difference between adult life expectancy to age 95 and adult life expectancy bounded by the age of the oldest cohort member was 0.12 years (0.04 years for men and 0.18 years for women); these constants were added to all estimates to account for survival beyond age 95.

We also investigated changes in the median length of life (i.e., the 50th percentile of the survival distribution) and in the full distribution of survival times. Pointwise 95% CIs on the survival curve were estimated (*29*). CIs for the median survival time were defined as the range of times for which the confidence bands on the survival function included the value 0.5. We constructed percentile-bootstrap CIs (1001 samples) for the difference in adult life expectancy between 2003 and 2011 and similarly for the difference in median length of life (*31*).

Finally, to illustrate the source of life expectancy gains, we quantified the change in age-specific mortality rates from 2003 to 2011, for 10-year age intervals, and estimated rate ratios using Poisson regression, with log-exposure time as the offset.

Between 1 January 2000 and 31 December 2011, 13,060 deaths occurred among 101,286 individuals aged 15 years and older, contributing a total of 651,350 person-years of follow-up time.

Adult life expectancy declined from 52.3 years in 2000 (95% CI, 50.9, 53.8) to 49.2 in 2003 (95% CI, 48.2, 50.3) (Fig. 1). These life expectancies are substantially lower than the 2000 World Health Organization (WHO) adult life expectancy estimates for South Africa as a whole (61.4 years), but they are similar to the estimates for neighboring Swaziland (54.6) and Lesotho (51.2) (*2*). There is substantial geographic variation in HIV rates within South Africa, and adult HIV prevalence in rural KwaZulu-Natal is more similar to that in Swaziland (23.6%) and Lesotho (24.5%) than to that in South Africa as a whole (17.1%) (*32*). From 2000 to 2003, adult life expectancy declined from 55.4 to 51.3 years for women and from 49.0 to 46.9 years for men (fig. S2).

In 2004, adult life expectancy started to increase, reaching 60.5 years in 2011 (95% CI, 59.0, 62.0), an 11.3-year gain (95% CI, 9.6, 12.9) in the mean length of life relative to 2003, the year before ART became available in the public-sector health system (Fig. 1). Both men and women experienced large gains in adult life expectancy over this period: 9.0 and 13.3 years, respectively (fig. S2). Sensitivity analyses using alternative definitions of the study population (e.g., excluding nonresident members of households located in the demographic surveillance area) yielded similar results (fig. S3). Annual estimates of adult life expectancy with 95% CIs for men, women, and both sexes are reported in table S2.

Comparing survival curves for 2003 and 2011 (Fig. 2), the median length of life rose from 42.6 years (95% CI, 41.2, 44.3) in 2003 to 60.7 years (95% CI, 58.8, 62.7) in 2011, an 18.1-year gain for a typical person in this population (95% CI, 15.4, 20.6). The change in the median is larger than the mean, because before 2004, the distribution of survival times was skewed to the right, so that the median length of life was less than the mean. Changes between 2003 and 2011 in the mean and median length of life were highly statistically significant ($P < 0.001$). Comparing the full distribution of lengths of life for 2003 and 2011 (Fig. 3) reveals a reduction in the proportion of deaths occurring in young adulthood.

Between 2003 and 2011, all-cause mortality declined by over 50% for adults aged 25 to 44. Mortality reductions at older ages were much smaller and not statistically significant (Table 1 and fig. S4). These changes in age-specific mortality rates are consistent with the decline in HIV-related mortality reported in previous studies of this cohort (*6, 7*).

As an extension of our analysis, we compared the observed changes in adult survival at the population level with the estimated costs of providing ART in this community between 2004 and 2011 to establish the cost-effectiveness of the past ART delivery. Our analysis complements previous studies that have used predictive models to project future costs and effects of ART (*33–35*).

Effects were assessed by comparing the total number of life years lived under the observed age-specific mortality rates between 2004 and 2011 with the number of life years that would have

Table 1. Age-specific mortality rates for 2003 and 2011.

Age	2003			2011			Rate ratio 2011 versus 2003		
	Deaths	PY*	Rate	Deaths	PY*	Rate	RR	95% CI	
15–24	121	204.7	0.6	67	218.5	0.3	0.52	0.39	0.70
25–34	387	131.4	2.9	201	160.2	1.3	0.43	0.36	0.51
35–44	308	84.7	3.6	143	86.5	1.7	0.45	0.37	0.55
45–54	175	53.3	3.3	117	59.6	2.0	0.60	0.47	0.75
55–64	108	31.1	3.5	90	34.2	2.6	0.76	0.57	1.00
65–74	109	21.5	5.1	84	20.1	4.2	0.83	0.62	1.10
75–84	72	8.9	8.1	89	11.5	7.7	0.95	0.70	1.30

*PY, hundreds of person-years. Age-specific mortality rates were estimated separately for 2003 and 2011. Rate ratios (RR) were estimated using a Poisson regression model, with log-exposure time as the offset. The proportion (number) of deaths due to HIV among persons aged 15 to 84 years was 59% (750) in 2003 and 46% (344) in 2010. At time of publication, verbal autopsy data were not yet complete for 2011.

Table 2. Life-year gains and program costs, 2004–2011.

Total life years gained, 2004–2011	8142
Estimated program costs, 2004–2011	$10,806,451
CER* ($ per life year)	$1593

*CER, cost-effectiveness ratio, defined as program costs per life year. Total life years gained is the difference between the number of adult life years lived between 2004 and 2011, based on observed mortality patterns and the number of adult life years that would have been lived had 2003 age-specific mortality rates persisted through 2011. Program costs were calculated by multiplying the total number of adult life years on ART or in pre-ART care by per-patient per-year cost estimates. All costs are reported as 2011 U.S. dollars. Total life years and program costs shown in the table are not discounted; the CER is based on cost and life-year estimates that were discounted at 3%. South Africa's per-capita GNI was $6960 in 2011; the CER, as a percentage of per-capita GNI, was 23%.

occurred during this period had the population been continuously exposed to mortality rates observed in 2003. Given that life expectancy probably would have continued to decline below 2003 levels in the absence of ART (Fig. 1), the 2003 mortality rates provide a conservative counterfactual. There were 436,135 life years lived between 2004 and 2011 based on observed mortality rates, and 427,993 life years in the counterfactual scenario without ART, a difference of 8142 life years (*23*) (Table 2 and table S3).

To estimate costs, we calculated the total number of person years on ART in the community in each year between 2004 and 2011 (fig. S1), and multiplied this by published costs of ART delivery for South Africa over this period, accounting for reductions in treatment costs over this period (*36*, *37*). Person years in pre-ART care were included at one-sixth of ART costs (*23*). During the period 2004–2011, we observed 8609 person years on ART and 7857 person years in pre-ART care (tables S3 and S4).

The total cost of ART in this population was estimated at $10.8 million over the study period. Discounting both costs and effects at 3%, the cost-effectiveness ratio (CER) was $1593 per life year saved, less than a quarter of South Africa's 2011 per-capita gross national income (GNI) (*38*) (Table 2). Interventions with CERs less than per-capita GNI, a standard lower bound on the monetary valuation of a life year, are considered very cost-effective (*35*). It is important to note that this high level of cost-effectiveness of ART delivery is achieved in a public-sector ART program in rural South Africa, where ART retention and adherence are imperfect and levels of treatment failure are high (*39*). Our study captures the full range of patient experiences on ART.

We describe here the full population-level impact of a public-sector ART program on adult life expectancy in a setting of high HIV prevalence in rural South Africa. Our estimates capture the net effects of ART scale-up on the survival of HIV patients receiving ART (direct effects), the mortality of people who are not on ART (spillover effects), and the unmasking of non–HIV-related mortality in HIV-infected people whose lives have been extended by ART (compositional effects). Although the reversal of the decline in adult life expectancy coincided with the scale-up of ART (Fig. 1 and fig. S1), our estimates may also capture mortality trends not linked to the scale-up of ART. First, other changes in the community may also have affected survival, such as rural electrification, improved access to safe water, expansion of non-HIV health services, or a growing burden of noncommunicable diseases. Second, HIV-specific mortality trends may be influenced by internal dynamics of the HIV epidemic; in particular, historical trends in HIV incidence.

To assess the contribution of non–HIV-related mortality to the observed gains in adult life expectancy, we estimated HIV-cause–deleted adult life expectancy (2001–2010) using verbal autopsy data collected in the population surveillance (*7*). Cause-deleted life expectancy provides a measure of the impact that a particular cause of death has on life expectancy. If cause-specific mortality risks are independent, then cause-deleted adult life expectancy provides an estimate of what adult life expectancy would be in the absence of HIV-related mortality (*29*), and a plausible upper bound on the life expectancy gains that could be attained from further investments in HIV treatment and prevention programs. HIV-cause–deleted adult life expectancy remained almost constant throughout the period between 2001 and 2010 at about 70 years, even as observed adult life expectancy increased from 49.2 years in 2003 to 58.7 years in 2010 (Fig. 4). These patterns imply a decline in HIV-related mortality rates amid stable mortality rates for other causes. This analysis confirms previous research in this population that found that secular changes in adult mortality between 2004 and 2009 were attributable to reductions in HIV-related mortality, with no systematic trends in mortality due to injuries, noncommunicable diseases, and other causes (*7*). The absence of any trend in HIV-cause–deleted life expectancy suggests that the overall changes in survival in the population during this period were not substantially driven by ART spillover effects, compositional effects, or changes in other mortality risks in the community.

The large reduction in HIV-related mortality after 2004 is consistent with direct effects of ART on the survival of HIV patients. However, these changes could be explained in part by historical patterns in HIV incidence. For example, if HIV infection rates peaked in the late 1990s, then a decline in mortality would be expected in the late 2000s, because of the 8- to 10-year latency period from HIV infection to death. We note, however, that HIV incidence has not declined in this area (*40*), and prevalence has increased (*12*); if anything, there are more people at risk for premature death due to HIV in recent years. One way to gauge the contribution of dynamics internal to the epidemic is to project trends in adult life expectancy in the absence of ART. Using the Actuarial Society of South Africa's 2008 AIDS Model, we predicted adult life expectancy for black South Africans from the beginning of the epidemic to 2011, under the assumption that ART was not available to adults (*21*). In contrast to the changes that we observed in individual-level surveillance data, the model projected a further decline in adult life expectancy between 2003 and 2011, in the absence of ART (fig. S5). Although we cannot rule out internal dynamics of the epidemic playing some role in the recovery of life expectancy, the widespread provision of ART through the public sector was almost certainly the most important factor explaining these changes.

We did not use disability or quality-of-life weights in our valuation of life-years lived with HIV on ART. Accounting for possibly lower quality of life would lower the estimated gains from additional years lived with HIV on ART. However, recent evidence suggests nearly complete recovery of physical and social functioning in people on ART (*41*), and the latest therapeutic regimens have reduced side effects (*42*), allowing people with HIV on ART to lead essentially normal lives. Further, our focus on life years, rather than disability-adjusted life years, allows us to include changes in health throughout the population, given that we do not observe non–HIV-related morbidity.

We observed gains of 11.3 years in adult life expectancy between 2003 and 2011, using individually measured data from a complete population cohort. These findings have several important implications. Our estimates suggest that existing predictions of changes in adult life expectancy based on demographic models rather than directly observed data, as in our study, have substantially underestimated the effects of ART scale-up on survival in HIV-hyperendemic populations. For example, using a modified two-parameter logit prediction model, WHO estimated that adult life expectancy in South Africa did not increase, but in fact declined from 61.4 years in 2000 to 58.7 years in 2009 and from 54.6 years to 53.4 years in neighboring Swaziland (*2*).

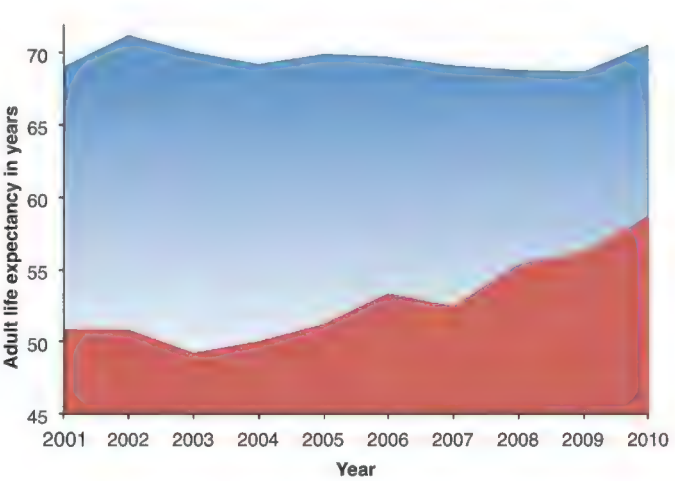


Fig. 4. HIV-cause–deleted adult life expectancy. Trends in adult life expectancy (red line) and HIV-cause–deleted adult life expectancy (blue line) for 2001–2010. HIV-cause–deleted life expectancy was estimated excluding deaths due to HIV, as identified by verbal autopsy in the Africa Centre surveillance. Whereas adult life expectancy increased after 2003, there was no systematic trend in HIV-cause–deleted adult life expectancy.

Additional gains in adult life expectancy for this population may be possible. In 2011, there was still substantial excess mortality due to HIV among younger adults under 50 years, as shown in Fig. 3. Increased efforts to recruit people with HIV into care and treatment earlier, to retain patients on treatment, and to ensure access to other health services may lead to further survival gains. Of particular interest in this setting is South Africa's 2011 expansion of treatment eligibility to all patients with CD4 < 350 cells/μl, which will facilitate earlier initiation on therapy. At present, only about half of those eligible for ART under the revised eligibility definition are receiving ART in South Africa (*43*). Although our findings strongly suggest that additional gains in life expectancy are possible, there are several sources of uncertainty regarding future trends. For one, although ART has been scaled up rapidly, sustaining and improving on existing survival gains will depend on continued political and financial commitment to ensuring access to treatment. Future mortality trends will also be influenced by the effects of ART scale-up on HIV acquisition (*26*), HIV prevalence, sexual behavior, care-seeking for HIV, and other health behaviors. Another important source of uncertainty is that the long-term survival of HIV patients on ART in this context is unknown, with treatment only widely available since 2004. However, evidence from this and other settings indicates that the risk of death for people with HIV actually declines with time after ART initiation (*13*).

The changes in adult life expectancy associated with ART scale-up in HIV-endemic populations are important information for governments and donors debating levels of support for public-sector HIV treatment programs. Changes in adult life expectancy resulting from ART may also have implications for forward-looking decisions of individuals, households, communities, and governments. In settings with high HIV prevalence (*12*) and high levels of social exposure to ART (*44*), we would expect individuals to revise their beliefs about their own longevity because of changes in survival in the community. These beliefs may influence, among other things, family planning, investments in human capital (such as schooling and job training), savings behavior, and willingness to engage in risks with negative consequences borne in the future (such as smoking, drug use, and criminal activity) (*45–48*). For households, communities, and countries, rising adult life expectancy will reduce the number of new orphans, improve the cross-generational transmission of knowledge and norms, and may lead to higher trust and social capital, as well as lower interest rates. For governments, rising adult life expectancy greatly increases the returns from investments in education and job training programs. Such changes will also have to be factored into projections of future pension obligations. Most important, gains in adult life expectancy provide the clearest evidence yet of the population-level impact of well-designed public-sector ART programs in settings of high HIV prevalence.

References and Notes

1. United Nations, Department of Economic and Social Affairs, Population Division, *World Population Prospects: The 2010 Revision* (CD-ROM Edition, 2011).
2. WHO, *World Health Statistics 2012. Life Tables for WHO Member States*; and methodological appendix *Life Tables for 2009, 2000 and 1990: Summary of Data and Methods Used* (WHO, Geneva, Switzerland, 2012).
3. P. Piot, *Science* **288**, 2176 (2000).
4. T. Barnett, A. Whiteside, *AIDS in the Twenty-First Century: Disease and Globalization* (Palgrave Macmillan, New York, 2002).
5. M. Egger *et al.*; ART Cohort Collaboration, *Lancet* **360**, 119 (2002).
6. A. J. Herbst *et al.*, *Bull. World Health Org.* **87**, 754 (2009).
7. A. J. Herbst, T. Mafojane, M. L. Newell, *Popul. Health Metr.* **9**, 47 (2011).
8. M. Nyirenda, V. Hosegood, T. Bärnighausen, M. L. Newell, *AIDS* **21** (suppl. 6), S73 (2007).
9. W. Muhwava, M. Nyirenda, *Demographic and Socioeconomic Trends in the ACDIS Africa Centre for Health and Population Studies* (Monograph Series No. 2, January 2008; www.africacentre.ac.za/Default.aspx?tabid=105).
10. A. Jahn *et al.*, *Lancet* **371**, 1603 (2008).
11. E. Bendavid, C. B. Holmes, J. Bhattacharya, G. Miller, *JAMA* **307**, 2060 (2012).
12. J. Zaidi, E. Grapsa, F. Tanser, M. L. Newell, T. Bärnighausen, "HIV prevalence trends after scale-up of antiretroviral treatment: A population-based study in a poor rural community in KwaZulu-Natal" (late-breaking abstract. International AIDS Conference, Washington, DC, 27 July 2012).
13. E. J. Mills *et al.*, *Ann. Intern. Med.* **155**, 209 (2011).
14. B. Žaba *et al.*, *Trop. Med. Int. Health* **17**, e3 (2012).
15. T. Bärnighausen *et al.*, *Lancet Infect. Dis.* **11**, 942 (2011).
16. D. Nash, Y. Wu, B. Elul, D. Hoos, W. El Sadr; International Center for AIDS Care and Treatment Programs, *AIDS* **25**, 1523 (2011).
17. T. Bärnighausen, D. E. Bloom, S. Humair, *Sex. Transm. Infect.* **88**, e2 (2012).
18. K. A. Grépin, *Health Aff.* **31**, 1406 (2012).
19. M. Mahy, J. Stover, K. Stanecki, R. Stoneburner, J. M. Tassie, *Sex. Transm. Infect.* **86** (suppl. 2), ii67 (2010).
20. Statistics South Africa, *Midyear Population Estimates. Statistical Release P0302*; and methodological appendix *A Methodology for Population Estimation at the National and Provincial Levels: The Approach Used by Statistics South Africa* (Republic of South Africa, Pretoria, South Africa, July 2011); www.statssa.gov.za/publications/statsdownload.asp?PPN=p0302&SCH=4986.
21. Actuarial Society of South Africa, *2008 AIDS and Demographic Model* (Cape Town, South Africa, March 2011).
22. F. Tanser *et al.*, *Int. J. Epidemiol.* **37**, 956 (2008).
23. Information on materials and methods, as well as supplementary tables and figures, are available on *Science* Online.
24. K. J. Rothman, S. Greenland, T. L. Lash, *Modern Epidemiology, 3rd Edition* (Lippincott Williams & Wilkins, Philadelphia, PA, 2008).
25. C. Day, P. Barron, N. Massyn, A. Padarath, R. English, Eds., *The District Health Barometer: 2010/2011* (Health Systems Trust, Durban, South Africa, 2011).
26. F. Tanser, T. Bärnighausen, E. Grapsa, J. Zaidi, M.-L. Newell, *Science* **339**, 966 (2013).
27. C. F. Houlihan *et al.*, *Int. J. Epidemiol.* **40**, 318 (2011).
28. J. P. Klein, M. L. Moeschberger, *Survival Analysis: Techniques for Censored and Truncated Data* (Springer, New York, ed. 2, 2003).
29. J. D. Kalbfleisch, R. L. Prentice, *The Statistical Analysis of Failure Time Data* (Wiley, New York, ed. 2, 2002).
30. E. L. Kaplan, P. Meier, *J. Am. Stat. Assoc.* **53**, 457 (1958).
31. B. Efron, R. J. Tibshirani, *Introduction to the Bootstrap. Monographs on Statistics and Applied Probability 57* (Chapman and Hall/CRC, Boca Raton, FL, 1993).
32. UNAIDS, *UNAIDS Report on the Global AIDS Epidemic* (Joint United Nations Programme on HIV/AIDS, Geneva, Switzerland, 2010).
33. T. Bärnighausen, D. E. Bloom, S. Humair, *Proc. Natl. Acad. Sci. U.S.A.* 10.1073/pnas.1209017110 (2012).
34. J. W. Eaton *et al.*, *PLoS Med.* **9**, e1001245 (2012).
35. S. J. Goldie *et al.*, *N. Engl. J. Med.* **355**, 1141 (2006).
36. S. Rosen, L. Long, I. Sanne., *Trop. Med. Int. Health* **13**, 1005 (2008).
37. K. Condliffe, "Facility-based unit costing for antiretroviral treatment in Ethiopia, Malawi, Rwanda, South Africa, and Zambia," paper presented at the Second International HIV Treatment as Prevention Workshop, Vancouver, Canada, 23 April 2012.
38. World DataBank, *World Development Indicators and Global Development Finance* (The World Bank Group, Washington, DC, 2012).
39. P. C. Mutevedzi *et al.*, *Bull. World Health Org.* **88**, 593 (2010).
40. T. Bärnighausen, F. Tanser, M. L. Newell, *AIDS Res. Hum. Retroviruses* **25**, 405 (2009).
41. J. Bor, F. Tanser, M. L. Newell, T. Bärnighausen, *Health Aff.* **31**, 1459 (2012).
42. Department of Health, *Clinical Guidelines for the Management of HIV & AIDS in Adults and Adolescents* (Republic of South Africa, Pretoria, South Africa, 2010); www.who.int/hiv/pub/guidelines/south_africa_art.pdf.
43. L. F. Johnson, *South. Afr. J. HIV Med.* **13**, 22 (2012).
44. J. Bor, T. Bärnighausen, C. Newell, F. Tanser, M. L. Newell, *Trop. Med. Int. Health* **16**, 988 (2011).
45. S. Jayachandran, A. Lleras-Muney, *Q. J. Econ.* **124**, 349 (2009).
46. J. Fortson, *Rev. Econ. Stat.* **93**, 1 (2011).
47. V. Baranov, H. P. Kohler, "The impact of AIDS treatment on savings and human capital investment in Malawi," paper presented at the Northeast Universities Development Consortium Conference, Dartmouth College, Hannover, NH, 3 November 2012.
48. D. E. Bloom, D. Canning, *Science* **287**, 1207, 1209 (2000).

Acknowledgments: We thank all the respondents who gave their time to this research and the staff of the Africa Centre for Health and Population Studies and the Hlabisa HIV Care and Treatment Programme. In particular, we thank all data collection and database management coordinators, supervisors, and managers (T. Mutevedzi, C. Newell, Z. Gqwede, H. Madida, L. Sithole, P. Dlamini, P. Gwala, T. Mngomezulu, N. Ntombela, N. Myeni, B. Ntimane), as well as all fieldworkers, verbal autopsy nurses, surveillance trackers, data capturers, document management clerks, and quality controllers. We also thank our funders: the Wellcome Trust (Africa Centre for Health and Population Studies); National Institutes of Health grants R01 HD058482-01 and 1R01MH083539-01 (T.B.); and the Harvard Global Health Institute and Harvard Center for Population and Development Studies (J.B.). Data are archived and accessible at no cost from the Africa Centre for Health and Population Studies. The authors report no conflicts of interest.

Supplementary Materials

www.sciencemag.org/cgi/content/full/339/6122/961/DC1
Materials and Methods
Figs. S1 to S5
Tables S1 to S4
References (*49–63*)

18 September 2012; accepted 1 February 2013
10.1126/science.1230413

High Coverage of ART Associated with Decline in Risk of HIV Acquisition in Rural KwaZulu-Natal, South Africa

Frank Tanser,[1]* Till Bärnighausen,[1,2] Erofili Grapsa,[1] Jaffer Zaidi,[1] Marie-Louise Newell[1,3]

The landmark HIV Prevention Trials Network (HPTN) 052 trial in HIV-discordant couples demonstrated unequivocally that treatment with antiretroviral therapy (ART) substantially lowers the probability of HIV transmission to the HIV-uninfected partner. However, it has been vigorously debated whether substantial population-level reductions in the rate of new HIV infections could be achieved in "real-world" sub-Saharan African settings where stable, cohabiting couples are often not the norm and where considerable operational challenges exist to the successful and sustainable delivery of treatment and care to large numbers of patients. We used data from one of Africa's largest population-based prospective cohort studies (in rural KwaZulu-Natal, South Africa) to follow up a total of 16,667 individuals who were HIV-uninfected at baseline, observing individual HIV seroconversions over the period 2004 to 2011. Holding other key HIV risk factors constant, individual HIV acquisition risk declined significantly with increasing ART coverage in the surrounding local community. For example, an HIV-uninfected individual living in a community with high ART coverage (30 to 40% of all HIV-infected individuals on ART) was 38% less likely to acquire HIV than someone living in a community where ART coverage was low (<10% of all HIV-infected individuals on ART).

One of the most successful public health interventions ever undertaken has been the provision of combination antiretroviral therapy (ART) to more than 6.2 million people in sub-Saharan Africa (*1*). The ART scale-up has resulted in substantial population-level reductions in HIV-related mortality in many populations (*2*, *3*) and overall is estimated to have saved a total of more than nine million life-years (*1*). The results of the landmark HIV Prevention Trials Network (HPTN) 052 trial in HIV-discordant couples demonstrated unequivocally that reducing the infected partner's viral load through ART substantially lowers the probability of HIV transmission to the uninfected partner (*4*). This finding has further fueled hope that widespread use of ART could not only substantially increase life expectancy but also reduce the rate of new HIV infections at a population level and reverse the epidemic (*5*). Indeed, predictive mathematical models have suggested that under certain conditions high coverage of ART could lead to a substantial decrease in the rate of new HIV infections (*6*, *7*).

The HPTN 052 trial was run under the controlled conditions of a well-conducted clinical study, and hopes of a substantial reduction in the rate of new HIV infections in the hyperendemic communities of sub-Saharan Africa have been tempered with legitimate concerns relating to uptake of HIV testing and treatment, retention, adherence, resistance development, risk compensation in sexual behavior, high rates of migration, and the capacity of health systems to deliver ART. Further, it is unclear to what extent the results of the trial can be extrapolated to communities where stable, cohabiting couples are not the norm (*8*).

The existing global evidence of the population effect of HIV treatment as prevention has been based on "ecological" associations (correlations between group-level variables and group-level outcomes) between increasing ART coverage and HIV outcomes, such as the number of new HIV diagnoses in a given administrative region of (*9*). Such designs provide a weak basis for causal inference and may be subject to "ecological fallacy" (*10*), that is, the inferential fallacy that occurs when a statistical association observed between variables on an aggregate level does not reflect the association that exists at an individual level. In this study, we use data from a population cohort of nearly 17,000 individuals who were HIV-uninfected at baseline and follow them up over several years (2004 to 2011), observing individual HIV seroconversions. We regress the individual-level outcome—time-to-HIV seroconversion—on ART coverage in the local community surrounding each HIV-uninfected individual to estimate the effect of increasing coverage on their risk of HIV acquisition. Although our exposure variable of interest—community-level ART coverage—is necessarily ecological in nature, our outcome and other variables are measured at the individual level, avoiding ecological fallacies in effect attribution. An alternative, rigorous approach to measure the impact of ART coverage on HIV acquisition risk is a cluster-randomized controlled trial. One such trial is already under way at the Africa Centre for Health and Population Studies, and several other such trials will be starting shortly (*11*). However, the results of these trials will not be available for at least another 4 years.

The study we report here took place in Hlabisa subdistrict, one of the five subdistricts in the rural district of Umkhanyakude in northern KwaZulu-Natal, South Africa (fig. S1). ART has been rapidly scaled up in the district through the Hlabisa HIV Treatment and Care Programme, and by mid-2012 treatment had been initiated in more than 20,000 patients (*12*). The study area is characterized by high adult HIV prevalence (24% in adults aged 15 years and older in 2011) and high levels of poverty and unemployment (in 2010, 67% of adults over the age of 18 in the study area were unemployed). ART was delivered in 17 community-based clinics by nurses and ART counselors using the standard South African drug regimens, which conform to World Health Organization (WHO) ART guidelines (*13*). Initially, adults with WHO clinical stage 4 disease or $CD4^+$ cell counts of <200 cells/μl were eligible for ART. The treatment eligibility threshold was increased to $CD4^+$ <350 cells/μl for pregnant women and tuberculosis patients in April 2010, and this new threshold was subsequently extended to all adults in August 2011. During the period of observation, no population groups were eligible for treatment initiation at $CD4^+$ ≥350 cells/μl.

Within the Hlabisa subdistrict lies the Africa Centre demographic surveillance area, with ~60,000 resident individuals under demographic surveillance at any given time (*14*). Between 2004 and 2006, all women age 15 to 49 years and men age 15 to 54 years resident in the surveillance area were eligible for HIV testing. In 2007, eligibility for the HIV survey was extended to cover all residents age ≥15 years. This HIV population cohort is open, and any individual who migrates into the area immediately becomes eligible for participation in the HIV survey. Within a 5-year period, about 80% of all individuals consented to HIV testing. However, in a single year of testing, the HIV consent rate was ~50%; the rate increased in the most recent 3-year period. To determine the proportion of the total HIV-infected population receiving ART, data from the Hlabisa HIV Treatment and Care Programme's database [maintained by Africa Centre personnel and described elsewhere (*12*)] were matched to the Africa Centre's demographic data using strict matching criteria: A person was linked either by his/her unique South African identification number or by his/her name, if both first and last name matched.

The 16,667 repeat testers who were HIV-uninfected at first test in the Africa Centre's population-based cohort were individually geolocated to their homesteads, using the comprehensive Africa Centre geographic information system. HIV prevalence and ART coverage for each individual's own surrounding community were then determined for every year of observation. ART coverage and HIV prevalence around each individual were measured by means of a moving two-dimensional Gaussian kernel of 3 km search radius (fig. S3) for each year of observation (2004 to 2011). The size of the kernel was determined from the results of previous work (*15*). The kernel moves systematically across the map and measures spatial variation in HIV prevalence [derived from the ~10,000 participants who consent to HIV testing each year (*14*)] across the surveillance area. It then uses the resulting prevalence and geographic distribution of the total eligible

[1]Africa Centre for Health and Population Studies, University of KwaZulu-Natal, South Africa. [2]Department of Global Health and Population, Harvard School of Public Health, Boston, MA, USA. [3]Centre for Paediatric Epidemiology and Biostatistics, University College London Institute of Child Health, London, UK.

*To whom correspondence should be addressed. E-mail: tanserf@africacentre.ac.za

population to calculate the number of HIV-infected individuals in the surrounding local community for each cell on the grid. In the same manner, each patient on ART was geolocated to their homestead of residence, and the number of individuals in the surrounding local community on ART was computed for each cell on the grid. These data are then used to derive ART coverage (which we define as the proportion of all HIV-infected individuals receiving ART) at every location. Given the scattered distribution of the population (homesteads are not concentrated into villages or compounds as in many parts of Africa), the method is well suited to this setting because it does not impose any static geographical boundaries on the data but uses the precise location of each individual (fig. S2) to derive a sensitive community-level measure that is both responsive to local variations and robust to the effects of random noise. The resulting ART coverage measure is a true population estimate because its denominator includes all HIV-infected individuals in this community, independently of whether they have ever accessed an ART program or not.

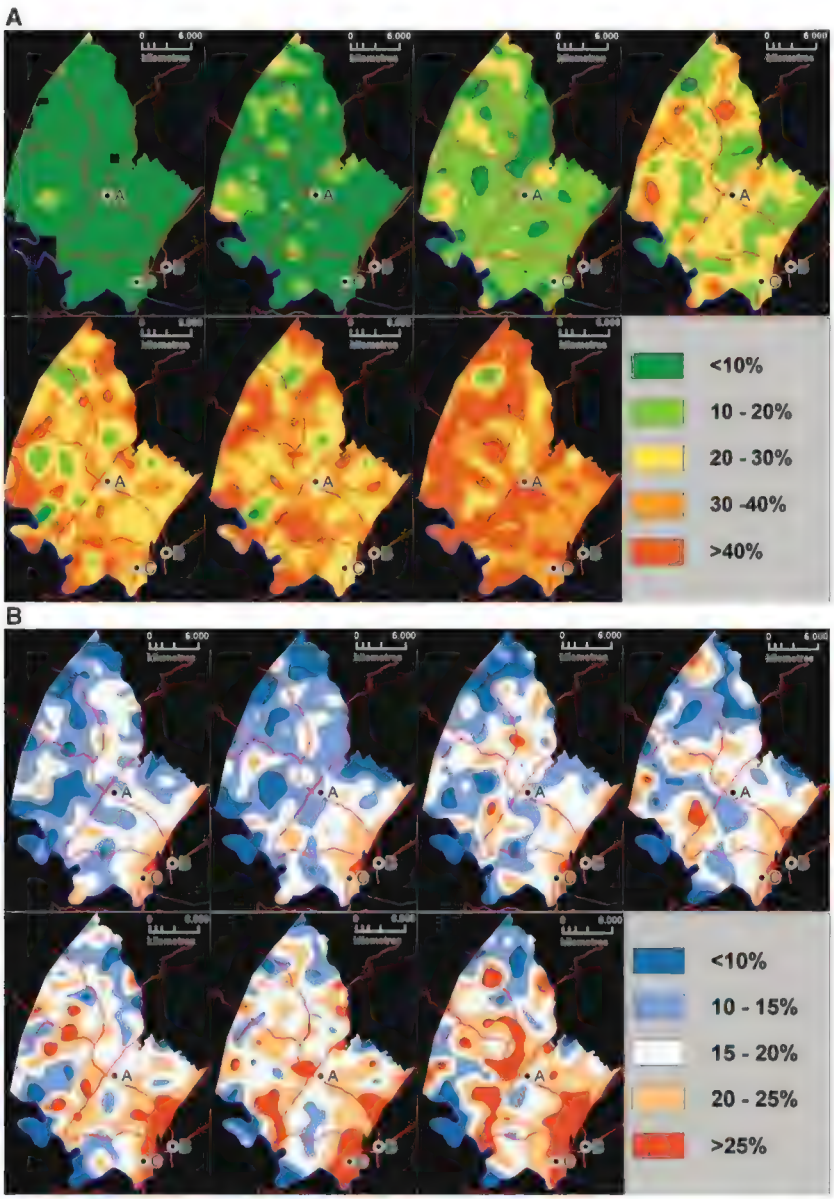


Fig. 1. Time series of maps showing the evolution of the proportion of the HIV-infected adults (≥15 years of age) receiving ART (**A**) and HIV prevalence (**B**) across the demographic surveillance area (2005 to 2008, left to right, top row; 2009 to 2011, left to right, bottom row). A, Africa Centre; B, Mtubatuba Town; C, KwaMsane Township. Main roads are also superimposed for ease of reference. Each pixel on the map corresponds to the proportion of the total HIV-infected population receiving ART (A) and total population infected with HIV (B) in the surrounding local community as measured using a standard Gaussian kernel of radius 3 km. Over the study period, estimated HIV prevalence in the adult population (≥15 years of age) increased from 18 to 24%, which is likely a consequence of the increasing life span of HIV-infected people due to ART. The proportion of the total HIV-infected population receiving ART was estimated at 37% in mid-2011.

The group of 16,667 repeat HIV testers in this study are 75% of all individuals (≥15 years of age) ever observed to be HIV-uninfected through the Africa Centre's population-based HIV surveillance (2004 to 2011). The median duration between tests in these individuals was 1.8 years. We used survival analysis to examine the effect of ART coverage in the surrounding local community on an HIV-uninfected individual's risk of HIV acquisition while controlling for a range of well-established demographic, behavioral, economic, and environmental determinants of acquisition of new HIV infection (*16*, *17*). These control variables rule out confounding by the individual and community characteristics they capture. In the analysis an individual from this group of repeat HIV testers was considered to be at risk of acquiring HIV after the date of their first HIV-seronegative test.

Treatment for the first patients in the surveillance area began in September 2004, and by July 2011, an estimated 37% of all HIV-infected adults in the area had been successfully started on ART. Overall, the speed with which ART has been scaled up in this population has been rapid but heterogeneous in time and space (Fig. 1A). In 2007, 6.4% of the population lived in communities where coverage of all HIV-infected individuals exceeded 30%, and by the middle of 2011 this figure had increased to 90.8%. The geography of ART coverage did not follow a clear spatial pattern over the period of observation (2004 to 2011). HIV prevalence was consistently highest in the peri-urban communities next to the National Road (Fig. 1B), where >30% of the adult population were infected with HIV. The lowest HIV prevalence occurred among communities in the more remote rural areas (<15%). There was little correlation between community-level HIV prevalence and ART coverage (correlation coefficient = 0.17).

We directly observed 1413 HIV seroconversions in the 53,605 person-years at risk observed over the period 2004 to 2011 in the 16,667 individuals who were HIV-uninfected at baseline. See table S1 for the characteristics of the HIV-uninfected individuals. The crude HIV incidence over the observation period was 2.63 new infections per 100 person-years [95% confidence interval (CI) 2.50 to 2.77]. Incidence peaked at 6.6 per 100 person-years in women at age 24, and 5 years later in men at 4.1 per 100 person-years of observation (Fig. 2).

After controlling for differences in the age and sex distribution, the risk of an HIV-uninfected individual acquiring HIV (that is, the acquisition hazard) was substantially lower in areas of high ART coverage (Fig. 3A). Specifically, the risk of infection to an individual living in an area where the ART coverage was 30 to 40% was 34% ($P < 0.0001$) less than to an individual living in an area with ART coverage of <10%. To increase the strength of causal inference based on the observed association between ART coverage and individuals' risk of newly acquiring HIV, we also controlled for a wide range of well-known demographic, behavioral, economic, and environmental determinants of an

individual's risk of HIV acquisition (community-level HIV prevalence, marital status, household-level wealth, and the number of sexual partners in the past 12 months). After ruling out confounding by these factors, there was a steep and highly significant decline in an individual's adjusted HIV acquisition hazard with increasing ART coverage (Fig. 3B). Holding other factors equal, an HIV-uninfected person who lived in a community with ART coverage of 30 to 40% was on average 38% ($P < 0.0001$) less likely to acquire HIV infection than an HIV-uninfected person living in a community with ART coverage of <10% (Fig. 3B). When we include ART coverage (%) in the model as a continuous variable we find that, overall, a 1% increase in ART coverage is associated with a 1.4% decline [adjusted hazard ratio (aHR) = 0.986] in the risk of acquisition of new HIV infection (Table 1).

As expected, a key independent predictor of the risk of acquiring new HIV infection was the level of existing HIV infection in the local community surrounding an HIV-uninfected individual (Fig. 3C). Controlling for other factors, an HIV-uninfected individual was 2.2 times as likely to acquire HIV in a community where HIV prevalence was >25% relative to the base category of <10% ($P < 0.0001$). In line with previous research conducted in this population, age, sex, number of sexual partners in the past 12 months, and marital status were other independent predictors of HIV acquisition (Table 1). The ART coverage hazard ratios were robust to the addition of different control variables and changed little across four nested models (table S2). Living in a peri-urban community was also associated with a greater risk of acquiring infection, but this effect was rendered insignificant by the addition of community-level HIV prevalence (table S2).

In the multivariable base-case model, we control for the known important determinants of HIV acquisition in this population (*16*, *17*). Moreover, the lack of spatial clustering observed in the model residuals (see supplementary materials) in this analysis suggests that unmeasured community-level factors are not influencing our results. Nevertheless, we also wanted to further exclude the potential influence of other HIV prevention services whose scale-up over space and time could have correlated with that of ART. To this end, we analyzed the time trends in six variables

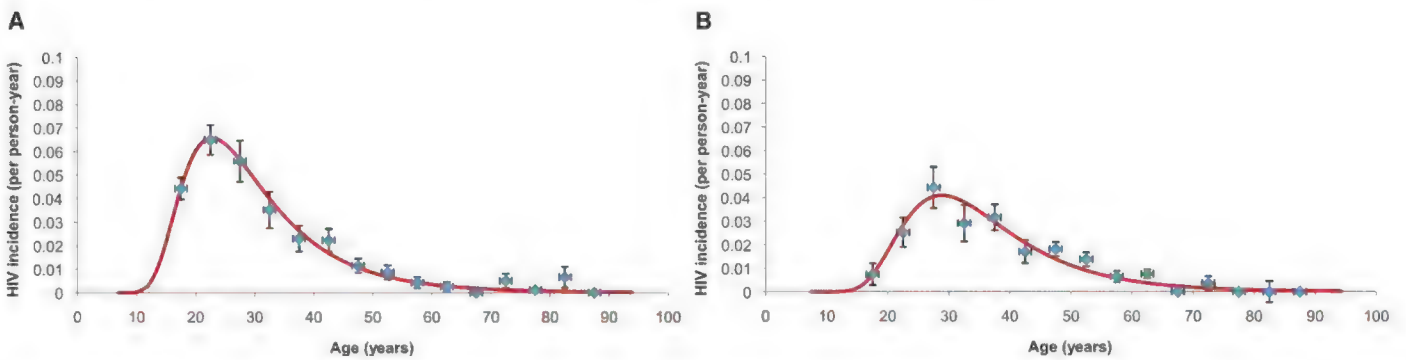


Fig. 2. Female (**A**) and male (**B**) age variations in HIV incidence (95% CI) by 5-year age-group for entire sample of repeat testers ($N = 16,667$; 53,605 person-years of observation). Superimposed on the graphs are log-normal functions (obtained by maximum likelihood) fitted to the incidence point estimates.

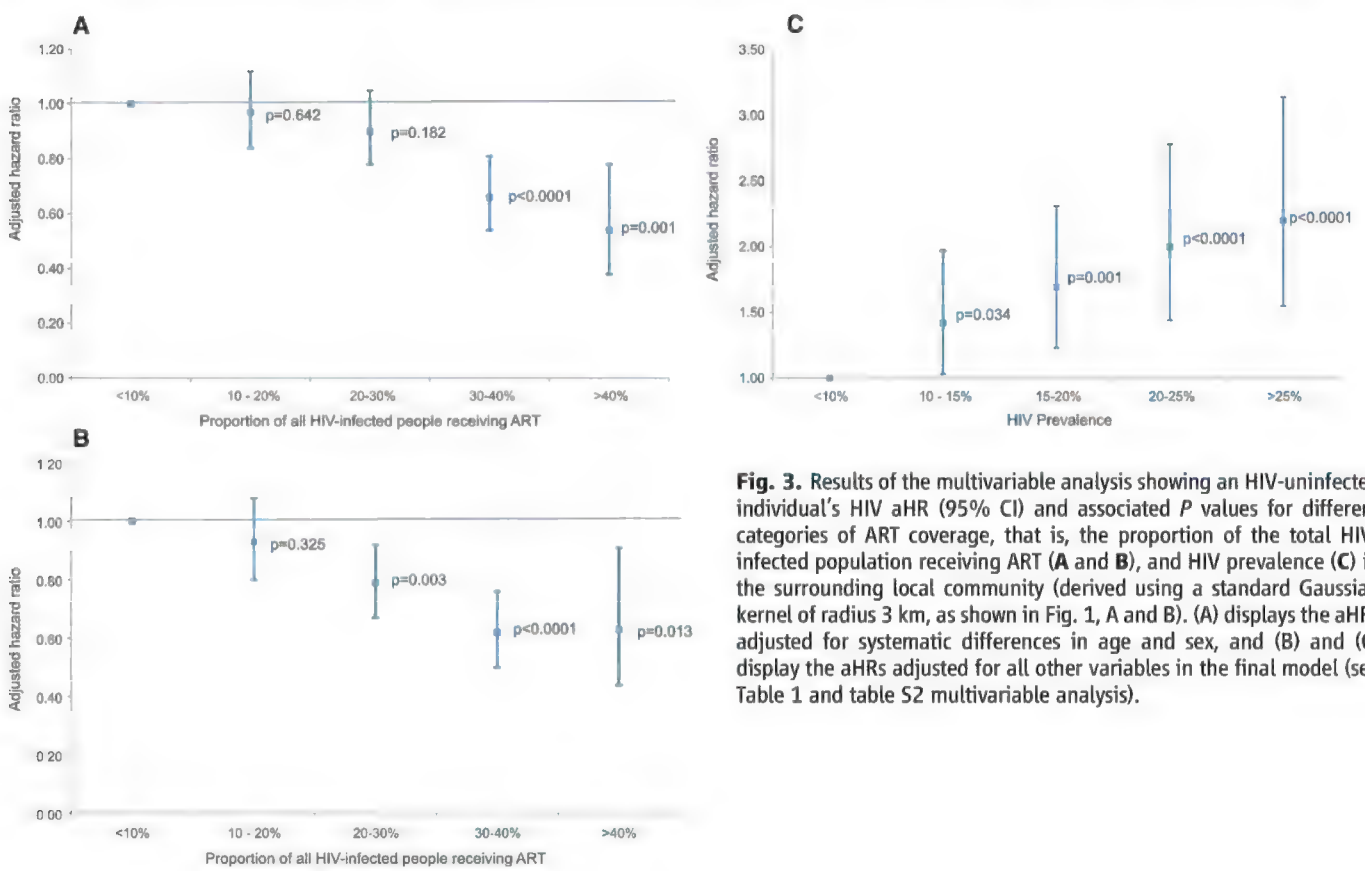


Fig. 3. Results of the multivariable analysis showing an HIV-uninfected individual's HIV aHR (95% CI) and associated *P* values for different categories of ART coverage, that is, the proportion of the total HIV-infected population receiving ART (**A** and **B**), and HIV prevalence (**C**) in the surrounding local community (derived using a standard Gaussian kernel of radius 3 km, as shown in Fig. 1, A and B). (A) displays the aHRs adjusted for systematic differences in age and sex, and (B) and (C) display the aHRs adjusted for all other variables in the final model (see Table 1 and table S2 multivariable analysis).

that capture the behavioral (*18*) or biological outcomes of other HIV prevention interventions: (i) the proportion of people who had ever had sex, (ii) the average number of sexual partners in the past year, (iii) the point prevalence of concurrent sexual partnerships, (iv) the mean age difference between partners, (v) proportion using condoms at last sexual activity, and (vi) proportion of circumcised men. We found that neither circumcision rates (which were 5.8% in 2004 and 5.0% in 2011) nor sexual partnership indicators (*18*) changed significantly over the study period. However, condom use at the last sexual activity with a regular partner increased significantly, from 26.2% in 2005 to 54.3% in 2011. We therefore analyzed the distribution of condom use across the surveillance area using the Gaussian kernel methodology (fig. S4) and ran a subanalysis in which we added condom use as a time- and space-varying community-level control variable to the multivariate base-case model. Controlling for condom use did not affect the strong relationship between the ART coverage and the risk of HIV acquisition (table S3 and fig. S5). This null finding is likely a consequence of the fact that condoms are only effective in preventing HIV infection if they are used correctly and consistently over a long period of time (*19*).

The adjusted hazard of acquiring HIV (fig. S7) shows a declining trend between 2004 and 2011. To investigate whether changes in the community over time, other than the behavioral and biological risk factors that we directly measured, might explain the ART coverage result, we added time-period dummy variables to our regression equation (table S4). After inclusion of these dummy variables, the ART coverage effect size estimates remain essentially unchanged.

The ART coverage result proved to be robust to a wide range of changes in the multiple sensitivity analyses we performed, including different survival model specifications (table S8), different age-eligibility criteria (table S10), different methods for imputing the date of HIV seroconversion (tables S5 and S6), and a range of different size

Table 1. Output from the final multivariable survival model showing the effect of coverage of ART in the surrounding community on an HIV-uninfected individual's hazard of acquiring HIV infection (N = 16,667).

Variable	Univariable analysis		Multivariable analysis†	
	HR (95% CI)	*P*	aHR (95% CI)	*P*
	Community-level ART coverage‡ (versus <10%)*			
10–20%	0.85 (0.74–0.98)	0.030	0.93 (0.8–1.08)	0.325
20–30%	0.72 (0.62–0.83)	<0.0001	0.79 (0.67–0.92)	0.003
30–40%	0.51 (0.42–0.63)	<0.0001	0.62 (0.50–0.76)	<0.0001
>40%	0.40 (0.28–0.58)	<0.0001	0.63 (0.44–0.91)	0.013
	Community-level HIV prevalence‡ (versus <10%)*			
10–15%	1.40 (1.02–1.93)	0.039	1.42 (1.03–1.97)	0.034
15–20%	1.63 (1.20–2.22)	0.002	1.69 (1.23–2.31)	0.001
20–25%	1.86 (1.37–2.54)	<0.0001	2.00 (1.44–2.78)	<0.0001
>25%	1.99 (1.45–2.74)	<0.0001	2.20 (1.55–3.14)	<0.0001
	Age-sex strata‡ (versus male 15–19 years old)			
Male 20–24	3.29 (2.34–4.64)	<0.0001	3.03 (2.15–4.27)	<0.0001
Male 25–29	5.77 (3.98–8.36)	<0.0001	5.22 (3.6–7.58)	<0.0001
Male 30–34	3.82 (2.35–6.2)	<0.0001	3.67 (2.23–6.04)	<0.0001
Male 35–39	4.15 (2.58–6.65)	<0.0001	4.32 (2.67–6.99)	<0.0001
Male 40–44	2.23 (1.23–4.05)	0.008	2.47 (1.36–4.5)	0.003
Male ≥ 45	1.25 (0.82–1.89)	0.294	1.71 (1.1–2.64)	0.017
Female 15–19	5.97 (4.40–8.09)	<0.0001	6.36 (4.67–8.64)	<0.0001
Female 20–24	8.45 (6.24–11.44)	<0.0001	8.92 (6.56–2.12)	<0.0001
Female 25–29	7.16 (5.13–9.99)	<0.0001	7.74 (5.52–0.85)	<0.0001
Female 30–34	4.62 (3.16–6.74)	<0.0001	5.38 (3.66–7.91)	<0.0001
Female 35–39	2.96 (1.99–4.39)	<0.0001	3.71 (2.46–5.58)	<0.0001
Female 40–44	2.91 (2.00–4.24)	<0.0001	3.61 (2.44–5.36)	<0.0001
Female ≥ 45	0.87 (0.60–1.24)	0.429	1.14 (0.78–1.67)	0.494
	Area of residence‡ (versus rural)			
Peri-urban	1.37 (1.23–1.53)	<0.0001	1.13 (0.98–1.3)	0.105
Urban	1.00 (0.70–1.43)	0.985	0.93 (0.64–1.35)	0.703
	Marital status (versus single)			
Married monogamous	0.33 (0.27–0.41)	<0.0001	0.67 (0.53–0.84)	0.001
Married polygamous	0.43 (0.25–0.75)	0.003	1.14 (0.65–2.01)	0.64
	Number of partners in the last 12 months (versus zero)			
One	3.22 (2.18–4.76)	<0.0001	1.77 (1.17–2.67)	0.008
More than one	3.72 (2.47–5.61)	<0.0001	2.49 (1.61–3.86)	<0.0001
	Household wealth quintile (versus poorest)			
2nd poorest	0.99 (0.84–1.16)	0.881	0.96 (0.82–1.14)	0.658
3rd poorest	1.03 (0.88–1.21)	0.721	0.97 (0.82–1.14)	0.676
4th poorest	1.02 (0.87–1.20)	0.813	0.92 (0.78–1.08)	0.296
Wealthiest	0.94 (0.78–1.12)	0.468	0.82 (0.68–1.00)	0.044
Weibull shape parameter	–	–	2.94 (2.80–3.13)	0.003

*Derived using a standard Gaussian kernel (radius 3 km) around each HIV-uninfected individual in the population cohort. †Corresponding values for a model in which community-level ART coverage and HIV prevalence are used as continuous variables (%); ART coverage aHR = 0.986 (95% CI 0.981 to 0.991), $P < 0.0001$; HIV prevalence aHR = 1.038 (95% CI 1.026 to 1.050), $P < 0.0001$. We sequentially added polynomial terms for the community-level HIV prevalence and ART coverage to this model, in order to allow greater flexibility in the functional form of the relationship between these factors and HIV acquisition. However, none of these polynomial terms were significant and we thus did not include them in the analyses shown here. ‡Indicates variables that vary with time.

(radius = 2.5 to 3.8 km) and shape kernels used in constructing the community-level variables (table S7). In the base-case analysis, values that were missing for two of the covariates, marital status and number of sexual partners in the past year, were imputed using multiple imputation (see supplementary materials). As a sensitivity analysis, we repeated our regressions with the sample of complete cases only, that is, the sample of individuals without any missing values. The ART coverage effect size estimates in this analysis remained very similar to those in the base-case analysis (table S9). As described above, the ART coverage results also persisted after we controlled for the influence of other prevention activities that had changed over the study period (table S3), as well as after adjusting for the trends in calendar time (table S4). Consistent with the "treatment-as-prevention" hypothesis, the decline in the risk of acquisition of HIV infection with increasing ART coverage in the older age groups (>35 years of age), where ART coverage was highest, was more pronounced than for the population as a whole (fig. S6). We also formally investigated whether selection effects could affect the results, by imputing HIV status for every individual who was eligible for testing in a particular year (see supplementary materials). We then used the same Gaussian kernel technique as for the base-case analysis to derive a prevalence map and resulting ART coverage estimates. Next, we reran the main model using these alternative community-level estimates. Using the imputed estimates to calculate community-level HIV prevalence and ART coverage had little bearing on the results (table S12). Further, to ascertain whether our findings would also continue to hold if we used an alternative (but less sensitive) approach, we ran an *isigodi* fixed-effects analysis, with HIV prevalence and ART coverage as time-varying variables. An *isigodi* is a traditional Zulu area (plural: *izigodi*); there are a total of 23 *izigodi* in this surveillance area. This analysis also confirmed a strong ART coverage effect (table S11).

Our findings build on the results from the recent HPTN 052 study that showed that if an HIV-uninfected person adheres to an effective antiretroviral therapy regimen, the risk of transmitting the virus to their uninfected sexual partner can be reduced by 96% (*3*). The trial was conducted under highly controlled conditions among stable, serodiscordant couples who had mutually disclosed HIV status to each other. High levels of ART adherence were achieved, and only 5% virologic failures in the intervention group were observed over the 42-month study period. By contrast, in the Hlabisa Treatment and Care Programme, 23% of patients initiated on ART had unsuppressed viral loads after having been on treatment for more than 12 months (*20*). The results presented are in line with recent modeling predictions in which an established stochastic mathematical model was calibrated specifically to the demographics, sexual behavior, and rate of ART scale-up observed in this population (*21*). The model projects a steady decline in the rate of new infections starting in 2010. Specifically, the model predicts an incidence rate ratio of 0.74 in 2011 (ART coverage at $CD4^+$ <350 cells/μl = 63%) decreasing to 0.57 in 2020 (ART coverage at $CD4^+$ <350 cells/μl = 84%) relative to the baseline year of 2004. Similarly, a systematic comparison of 12 mathematical HIV epidemic models demonstrated consistency in the prediction that high levels of ART coverage will reduce the rate of new infections (*6*).

In contrast to previous work that focused on ecological associations between uptake of ART and aggregate outcomes associated with the rate of new HIV infections in a given geographical unit (*22*–*29*), we directly measured time to HIV seroconversion in each individual and were thus able to avoid ecological fallacies and to control for a wide range of potential individual-level confounders in the analysis. By including both community-level ART coverage and HIV prevalence in the model as space- and time-varying variables, we have accounted for (and exploited for effect size estimation) the rapid scale-up in ART and associated changes in HIV prevalence in this population between 2004 and 2011 and the fact that each HIV-uninfected individual in the population will have differing exposures over time and space. It is a further strength of our data and approach that our effect size estimates capture the combined effect of multiple pathways through which ART coverage might reduce the risk of new HIV infection that are in addition to the biological effect of reducing mean viral load and consequently HIV transmission probability. For instance, HIV counseling and testing uptake might conceivably have increased because the availability of ART adds an important motivation to the other reasons for knowing one's HIV status. The increasing exposure to HIV counseling, in turn, could have led to changes in sexual behavior in both HIV-uninfected and HIV-infected individuals. Other examples of ART-induced behavior changes include risk compensation (*30*) (because the availability of ART mitigates the consequences of HIV infection) and decreased sexual risk-taking (because of increased optimism about the future as ART increases life expectancy) (*3*). Such behavioral effects of ART may contribute to the total effect of ART coverage on HIV acquisition and would have been captured by our approach. Finally, because ART is delivered through a public-sector program in a typical rural sub-Saharan African setting, we were able to test the effectiveness of ART in reducing the risk of acquiring infection in the context of a successful, but imperfect, real-world ART program (*3*).

A potential threat to the validity of our finding is the possibility that individuals within low ART-coverage communities differ systematically from their counterparts in high ART-coverage communities in factors that affect HIV acquisition. Thus, we exploited the detailed longitudinal, individual-level information available in the Africa Centre cohort to control for the well-known determinants of HIV acquisition in this population, as well as for the potential effects of other synchronous HIV prevention interventions. However, as in any large population-based prospective cohort study, such as the Framingham Heart Study, we cannot completely rule out biases due to unobserved confounding factors. Furthermore, it is possible that in other populations, the relationship between ART coverage and HIV acquisition could differ—for instance, due to different sexual behaviors. Future studies, in particular cluster-randomized controlled trials, are necessary to confirm and test the validity and generalizability of our findings.

Another issue to consider is the potential effect of "contamination" arising from differences in the geography of sexual partner choice. In this regard, it is important to note that our ART coverage results were robust to a range of different size and shape "communities" (kernels) used in the analysis (see supplementary materials). In addition, we have previously shown that sexual partner choice in this population has a strong local geographical dimension (*17*). Indeed, over the duration of this study, 67% of individuals in the cohort report having at least one sexual relationship with someone in the same small Zulu community or *isigodi*. (*Izigodi* have a median area of 16.9 km^2, just 37% of the area searched around each HIV-uninfected participant to derive the community-level variables using the largest Gaussian kernel in our analyses.) This observation is further supported by the striking relationship between local HIV prevalence and risk of new infection (Fig. 3C). Moreover, such unsystematic contamination effects, whereby individuals choose some sex partners outside their community, will likely lead to attenuation of the relationship between the exposure and the outcome, biasing the results toward the null hypothesis (*31*). It is thus unlikely that we have overestimated the effect of ART coverage on HIV acquisition due to contamination.

Our results suggest that population-level reductions in the transmission of HIV can be achieved in nurse-led, devolved, public-sector ART programs in rural sub-Saharan African settings where complete coverage of therapy under existing treatment guidelines has not yet been attained. Although key questions remain about durability of protection, long-term adherence, and transmission of ART-resistant strains to partners (*32*), our findings indicate that commitment to further increasing ART coverage in sub-Saharan Africa can contribute substantially to meeting the Political Declaration of the United Nations General Assembly (*33*) of halving sexual transmission of HIV by 2015.

References and Notes

1. UNAIDS, "Together we will end AIDS" (UNAIDS, Geneva, 2012).
2. A. J. Herbst *et al.*, *Bull. World Health Organ.* **87**, 754 (2009).
3. J. Bor, A. J. Herbst, M.-L. Newell, T. Bärnighausen, *Science* **339**, 961 (2013).
4. M. S. Cohen *et al.*; HPTN 052 Study Team, *N. Engl. J. Med.* **365**, 493 (2011).
5. R. M. Granich, C. F. Gilks, C. Dye, K. M. De Cock, B. G. Williams, *Lancet* **373**, 48 (2009).
6. J. W. Eaton *et al.*, *PLoS Med.* **9**, e1001245 (2012).
7. R. Granich *et al.*, *PLoS ONE* **7**, e30216 (2012).
8. J. D. Shelton, *Science* **334**, 1645 (2011).

9. K. Smith, K. A. Powers, A. D. M. Kashuba, M. S. Cohen, *Current Opinion in HIV and AIDS* **6**, 315 (2011).
10. M. Susser, *Am. J. Epidemiol.* **133**, 635 (1991).
11. R. Granich *et al.*; ART in Prevention of HIV and TB Research Writing Group, *Curr. HIV Res.* **9**, 446 (2011).
12. C. F. Houlihan *et al.*, *Int. J. Epidemiol.* **40**, 318 (2011).
13. South African National Department of Health, "National Antiretroviral Treatment Guidelines" (National Department of Health, Pretoria, South Africa, 2010).
14. F. Tanser *et al.*, *Int. J. Epidemiol.* **37**, 956 (2008).
15. F. Tanser, T. Bärnighausen, G. S. Cooke, M. L. Newell, *Int. J. Epidemiol.* **38**, 1008 (2009).
16. T. Bärnighausen *et al.*, *AIDS* **22**, 139 (2008).
17. F. Tanser *et al.*, *Lancet* **378**, 247 (2011).
18. N. McGrath, J. W. Eaton, F. Tanser, T. Bärnighausen, M.-L. Newell, Sexual behaviour trends by gender in a rural South African population-based cohort during the era of scaled-up access to VCT and ART, 2005-2010. Paper presented at the XIX International AIDS Conference, Washington, D.C., 22 to 27 July 2012.
19. S. Weller, K. Davis, *Cochrane Database Syst. Rev.* **1**, CD003255 (2002).
20. P. C. Mutevedzi *et al.*, *Bull. World Health Organ.* **88**, 593 (2010).
21. J. A. Hontelez *et al.*, *PLoS ONE* **6**, e21919 (2011).
22. A. Castel *et al.*, Use of community viral load as population-based biomarker of HIV, Washington, DC, 2004–2008. Paper presented at the 18th Annual Conference on Retroviruses and Opportunistic Infections, Boston, MA, 2011.
23. M. Das *et al.*, *PLoS ONE* **5**, e11068 (2010).
24. J. S. Montaner *et al.*, *Lancet* **376**, 532 (2010).
25. E. Wood *et al.*, *BMJ* **338**, b1649 (2009).
26. C. T. Fang *et al.*; Division of AIDS and STD, Center for Disease Control, Department of Health, Executive Yuan, *J. Infect. Dis.* **190**, 879 (2004).
27. M. H. Katz *et al.*, *Am. J. Public Health* **92**, 388 (2002).
28. M. G. Law *et al.*; Australian HIV Observational Database (AHOD), *J. Int. AIDS Soc.* **14**, 10 (2011).
29. T. C. Porco *et al.*, *AIDS* **18**, 81 (2004).
30. K. K. Venkatesh, T. P. Flanigan, K. H. Mayer, *AIDS* **25**, 1939 (2011).
31. E. Budtz-Jørgensen, N. Keiding, P. Grandjean, P. Weihe, R. F. White, *Stat. Med.* **22**, 3089 (2003).
32. A. Anglemyer, G. W. Rutherford, R. C. Baggaley, M. Egger, N. Siegfried, *Cochrane Database Syst. Rev.* **2011**, CD009153 (2011).
33. United Nations General Assembly, "Political declaration on HIV/AIDS: Intensifying our efforts to eliminate HIV/AIDS" (United Nations General Assembly, New York, 2011).

Acknowledgments: Supported by grant 1R01-HD058482-01 from the National Institute of Child Health and Human Development. Funding for the Africa Centre's Demographic Surveillance Information System and Population-Based HIV Survey was received from the Wellcome Trust, UK (grant 082384/Z/07/Z).

Supplementary Materials

www.sciencemag.org/cgi/content/full/339/6122/966/DC1
Materials and Methods
Figs. S1 to S7
Tables S1 to S12
References

30 July 2012; accepted 22 January 2013
10.1126/science.1228160

Minimal "Self" Peptides That Inhibit Phagocytic Clearance and Enhance Delivery of Nanoparticles

Pia L. Rodriguez,[1] Takamasa Harada,[1] David A. Christian,[1] Diego A. Pantano,[1] Richard K. Tsai,[1] Dennis E. Discher[1,2]*

Foreign particles and cells are cleared from the body by phagocytes that must also recognize and avoid clearance of "self" cells. The membrane protein CD47 is reportedly a "marker of self" in mice that impedes phagocytosis of self by signaling through the phagocyte receptor CD172a. Minimal "Self" peptides were computationally designed from human CD47 and then synthesized and attached to virus-size particles for intravenous injection into mice that express a CD172a variant compatible with hCD47. Self peptides delay macrophage-mediated clearance of nanoparticles, which promotes persistent circulation that enhances dye and drug delivery to tumors. Self-peptide affinity for CD172a is near the optimum measured for human CD172a variants, and Self peptide also potently inhibits nanoparticle uptake mediated by the contractile cytoskeleton. The reductionist approach reveals the importance of human Self peptides and their utility in enhancing drug delivery and imaging.

Macrophages evolved to engulf and clear invading microbes and dying cells, but they respond similarly to injected particles, viruses, and implants, which hampers delivery of therapeutics and imaging agents. Coating of nanoparticles or liposomes with polyethylene glycol (PEG) creates "stealth" brushes that mimic a cell's glycocalyx and that delay immune clearance of foreign particles (*1*–*3*), but brushes can also hinder uptake by diseased cells (*4*). Neither a polymer brush nor a glycocalyx stops adsorption of abundant serum proteins, such as immunoglobulin G (IgG) (table S1), which promote clearance [e.g., (*1*–*3*, *5*)], and any foreign polymer can also be immunogenic (*6*). Targeting diseased cells with ligand-modified particle constructs might make brushes unnecessary, but some ligands also promote rapid clearance by phagocytes (*2*, *7*). In the reductionist approach here, we make a synthetic, human-based "Self" peptide that specifically binds and signals to phagocytes to inhibit clearance of particles as small as viruses.

CD47 glycoprotein is a putative "marker of self" in mice (*8*) and is expressed on all cell membranes in humans, mice, and other mammals (*9*). It associates in cis with integrins (*10*) and other species-specific, immunogenic complexes on cells (*11*–*13*). Mouse knockouts of CD47 (mCD47) are viable, but when red blood cells (mRBCs) from these mice are injected into the circulation of control mice, the deficient cells are cleared within hours by splenic macrophages, whereas normal mRBCs circulate for weeks (*8*). CD47's extracellular domain interacts with CD172a (also known as signal regulatory protein-α, SIRPα) on phagocytes (*10*). Although binding is typically species-restricted (*13*), SIRPα is highly polymorphic, even within a species (*14*). Macrophages in nonobese diabetic/severe combined immunodeficient (NOD/SCID) strains of mice express a SIRPα variant that cross-reacts with human CD47 (hCD47), which explains why human hematopoietic cells engraft and circulate in NOD/SCID better than any other mice (*14*, *15*). In vitro, the CD47-SIRPα interaction inhibits mouse macrophage uptake of antibody-coated mRBCs (*8*), as well as human macrophage uptake of both human RBCs (hRBCs) and hCD47-coated microparticles (*16*). This is not surprising as SIRPα signaling inhibits contractility-driven uptake of micron-size cells and particles (*16*). However, contractile forces are widely considered unimportant to internalization of nanoparticles and viruses, and so it is unclear whether this inhibitory interaction could be exploited in nanoparticle-based therapeutics.

We addressed whether hCD47 and a synthetic Self peptide can minimize phagocytic uptake of nanoparticles and thereby enhance delivery in NOD/SCID mice with X-linked severe combined immunodeficiency (*Il2rg*$^{-/-}$) mice (NSG). We first showed that blocking mCD47 accelerates clearance of mRBCs in NSG mice. Cells (or nanobeads) were split into two samples, with one sample labeled by red fluorophore and the other sample labeled by far-red fluorophore plus antibody against mCD47. The samples were mixed 1:1 for injection into the same mouse, and blood samples at subsequent time points were analyzed by flow cytometry for both colors, which produced a ratio (fig. S1A) that minimizes mouse-to-mouse variations. IgG and other serum proteins physisorb in vivo to RBCs (*17*), viruses (*18*), and PEGylated nanoparticles (*19*) (table S1), but in NSG mice, IgG is very low or not detectable (versus ~100 μM in normal animals). Controlled opsonization with IgG was therefore used in most of our studies to better mimic immune-competent animals. With RBCs, mRBC-specific antibody was added to promote clearance via phagocytosis (*20*). Consistent with a marker-of-self function of mCD47, the persistence ratio for the mixed sample [mRBCs/(mRBCs with blocked mCD47)] increased exponentially with a doubling time (T) of 33 min (Fig. 1B); single-color analyses also give $T = 30$ min (fig. S1B).

[1]Molecular and Cell Biophysics and NanoBioPolymers Laboratory, University of Pennsylvania, Philadelphia, PA 19104, USA. [2]Pharmacological Sciences Graduate Group, University of Pennsylvania, Philadelphia, PA 19104, USA.

*To whom correspondence should be addressed. E-mail: discher@seas.upenn.edu

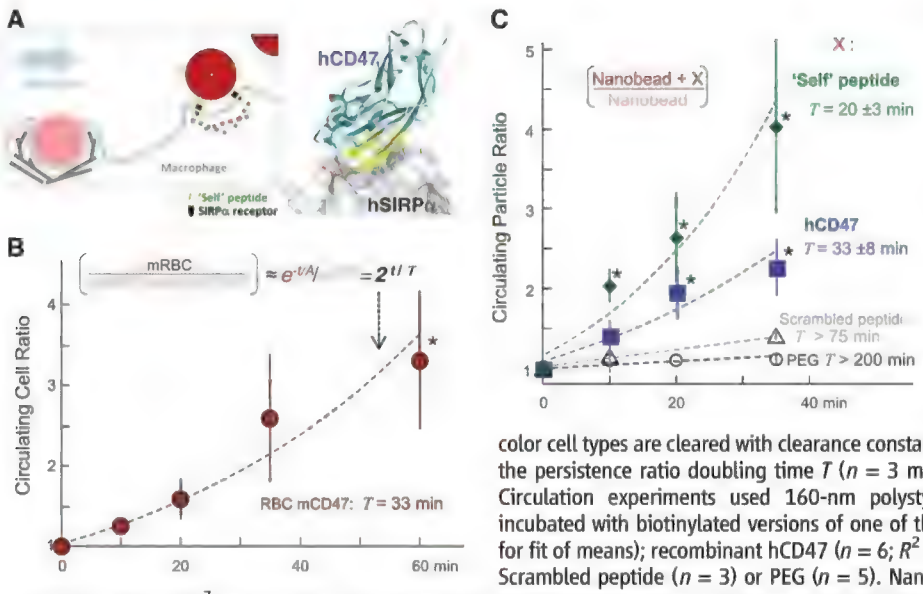


Fig. 1. Self peptide and hCD47 prolong the circulation of nanobeads in NSG mice. (**A**) Competitive circulation in which two colors of nanobeads or cells injected into the same mouse are flowing with blood and being cleared by a splenic macrophage (left) or else recognized as self and released (right). Blood was sampled periodically in the experiment (50 μl), and flow cytometry analysis yields the bead or cell ratio in each mouse at each time point. The cocrystal structure shows hCD47 with the Self peptide binding hSIRPα. (**B**) Competitive circulation experiment in which mRBCs from NSG mice were either blocked with anti-CD47 or not and were also opsonized with excess mRBC-specific antibody before cells were mixed together and injected into the tail vein. Both color cell types are cleared with clearance constants ~*A* or *B* (*A* > *B*), and the ratio of exponentials gives the persistence ratio doubling time *T* (n = 3 mice; R^2 = 0.93 for fit of means with indicated *T*). (**C**) Circulation experiments used 160-nm polystyrene beads with covalently attached streptavidin incubated with biotinylated versions of one of the following: synthetic Self peptide (n = 4; R^2 = 0.94 for fit of means); recombinant hCD47 (n = 6; R^2 = 0.92 for fit of means); or negative controls of either Scrambled peptide (n = 3) or PEG (n = 5). Nanobeads were also opsonized with streptavidin-specific antibody, and then 10^7 were injected. Flow cytometry quantification was typically done on 100 to 10,000 particles at each time point and typically included quantification of both hCD47 and opsonin on the nanobeads. For hCD47 and Self peptide, a separate fit for each mouse gives the indicated mean *T* ± SEM for each group, which is within 10% of the *T* obtained from fitting the group averages (dashed curves). Most data points for hCD47 and Self peptide differ significantly from PEG-nanobeads (*P < 0.05). All data are means ± SEM.

RBC membranes have hundreds of different interacting proteins, and many involved in clearance are different for mouse and humans (*12*, *21*). To give a better-defined surface for reductionist studies in vivo and also to begin testing the marker-of-Self concept on foreign particles of potential use for imaging and therapy, the extracellular immunoglobulin-like domain of hCD47, which binds SIRPα, was recombinantly expressed; site-specifically biotinylated; and then bound to streptavidin-coated, 160-nm polystyrene nanobeads. Beads were also labeled with red or near-infrared dyes (or left unlabeled) and controllably opsonized with either antibody against streptavidin (fig. S1C) or a biotinylated antibody for targeting (fig. S1D). After injection into an NSG mouse, blood analysis by flow cytometry clearly identified nanobeads on the basis of both distinctive scatter and fluorescence detection of the opsonizing antibody ± hCD47 (Fig. 1A and fig. S1, A and C). A persistence ratio for [(nanobead + hCD47)/nanobead] was well-controlled at every time point and again increased exponentially with a doubling time (*T*) of 33 ± 8 min (Fig. 1C). Mice injected with a single color of nanobead gave similar results (*T* = 31 min) (fig. S1E). PEG-biotin nanobeads that were also preopsonized showed a flat persistence curve (*T* > 200 min), consistent with the fact that PEG brushes alone do not directly inhibit clearance by macrophages (fig. S2, A to C). In the absence of preopsonization, PEG-nanobeads did circulate for hours as expected, but hCD47, once again, enhanced circulation (fig. S2D). hCD47 on virus-size particles is thus an inhibitor of in vivo clearance and thereby prolongs circulation.

Minimization of the 117–amino acid immunoglobulin-like domain of hCD47 to a small, binding-site Self peptide could provide key evidence that signaling to mouse SIRPα (mSIRPα) is part of the molecular mechanism for inhibiting clearance in vivo. A crystal structure of hCD47-hSIRPα suggests three distinct binding sites (*22*), but the highest density of interactions are in one loop in hCD47 between canonical β strands F and G, where a nine–amino acid sequence constitutes 40% of hCD47's contacting residues (Fig. 1A, structure). We designed by simulation a 21–amino acid Self peptide around this sequence with the aims of minimizing species specificity (*13*) and eliminating glycosylation of CD47, which impedes binding (*23*). Biotinylation on an amino-terminal PEG linker provided a means of attachment to streptavidin beads for in vivo studies. Relative to control nanobeads, the Self peptide increased the persistence of beads in the circulation with a mean doubling time, *T*, of 20 ± 3 min, whereas Scrambled peptide (see supplementary materials) had little impact on circulation (Fig. 1C). An apparent difference in persistence in the circulation between Self peptide and hCD47 is not significant (P = 0.18).

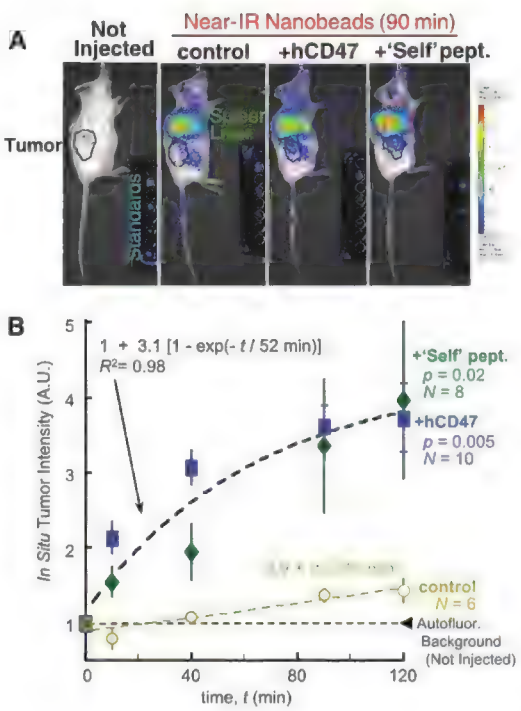


Fig. 2. Self peptide and hCD47 enhance tumor imaging by near-infrared particles. (**A**) NSG mice with flank tumors of A549 lung–derived cells (black circles) received tail vein injections of nanobead mixtures in which one bead type is labeled with DiR fluorophore. Images of live mice and calibration standards were taken with a Xenogen imager. Tumor-bearing mice have persistence ratios of particles in blood at 35 min, similar to results in Fig. 1C, even though many particles are seen in spleen and liver. (**B**) The tumor was located by bright-field imaging, and total fluorescence was quantified at each time point. All results for Self peptide and hCD47 were combined in the fit. *N*, Number of tumors from three different sets of tumor-bearing mice. All data are means ± SEM.

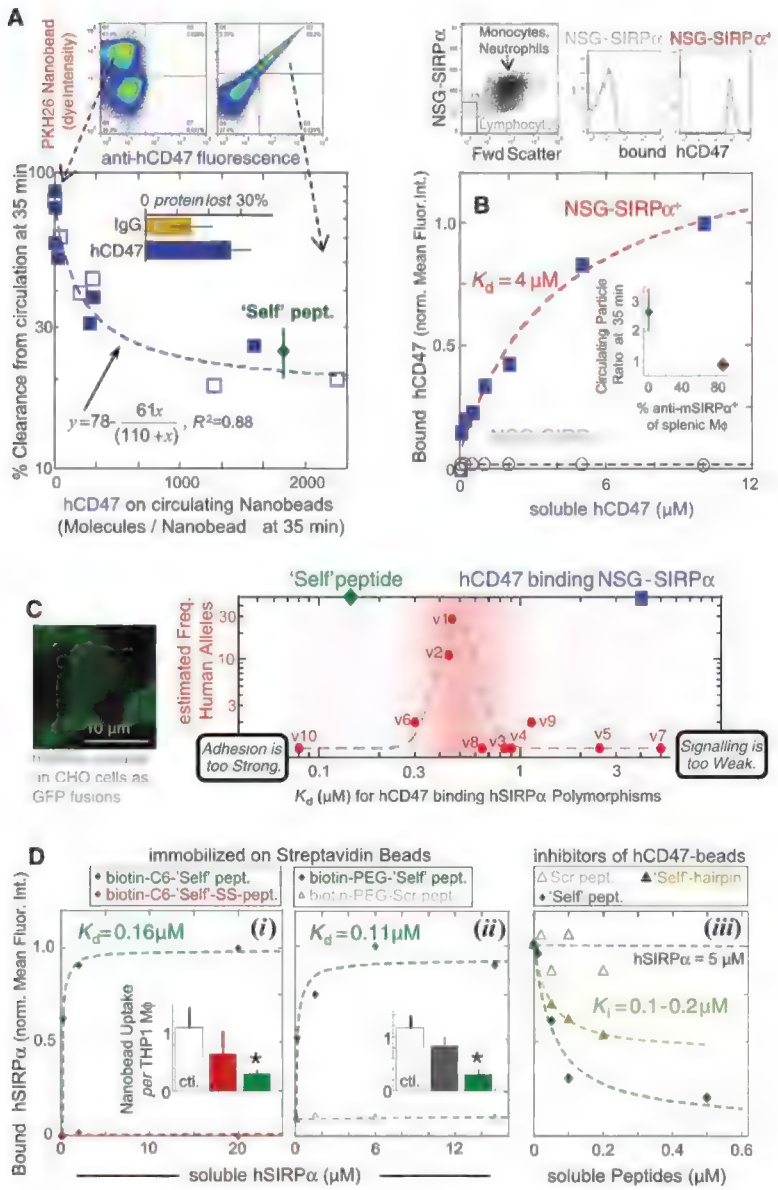


Fig. 3. Persistence of hCD47 and Self nanobeads depends on hCD47 density, consistent with low-affinity binding to NSG mSIRPα relative to hSIRPα variants. (**A**) The number of hCD47 molecules on the 160-nm beads 35 min after injection was either measured in two-color experiments (solid symbols; $n = 7$ mice) or single-color experiments (open symbols; $n = 6$ mice), with an average of 30% protein lost in circulation (inset). Self peptide levels are estimated to have a similar loss ($n = 4$ mice). Fluorescent nanobeads ($PKH26^+$ in flow cytometry, top) were confirmed by forward and/or side scatter, and fluorescent hCD47-specific antibody measured hCD47 levels (left, control nanobead sample; right, hCD47 nanobead sample). The inhibition curve gives $K_i = 110$ molecules/nanobead. (**B**) Affinity of soluble hCD47 for NSG neutrophils and monocytes, from flow cytometry analysis of Cy5-biotin–specific antibody. Lymphocytes are negative for SIRPα and do not bind soluble hCD47. (Inset) Preinjection of blocking anti-mSIRPα eliminates the enhanced circulation of Self nanobeads of Fig. 1C. After 35 min, most splenic macrophages (Mϕ: F4/80 antibody+) have mSIRPα antibody on their surface ($n = 4$ mice). (**C**) Ten reported variants of hSIRPα's N-terminal domain (*14*) were displayed on CHO cells to determine effective K_d values for soluble hCD47; soluble hSIRPα binding to hCD47 beads showed the same trend (fig. S7E). The putative allele frequency (*14*) is highest at intermediate K_d, whereas the affinity of hCD47 for mSIRPα on NSG phagocytes (blue square) is weaker and that of Self peptide for hSIRPα(v1) is stronger (green diamond). The Lorentzian fit is inspired by other mechanobiological signaling processes and has the form: $y = 1 + 0.05x^{11}/(0.50^{11} + x^{11})^2$, $R^2 = 0.85$. (**D**) Binding of peptides on beads to soluble hSIRPα (v1) was assayed by flow cytometry. Neither the Self-SS peptide with a T107C substitution nor the Scrambled peptide bind hSIRPα. The assays in (*iii*) use soluble peptides and show the 10–amino acid Self hairpin centered on the loop is a partial inhibitor. Bar graph insets in (*i* and *ii*) show in vitro phagocytosis assay results with the human THP1 cell line, which demonstrates that only the Self peptide (attached to biotin via either PEG or C6, 6-aminohexanoic acid) significantly inhibits phagocytic uptake (*$P < 0.05$ different from control). All data are means ± SEM.

Prolonged circulation of hCD47 beads and Self beads is based on a delay of phagocytic clearance by the spleen and perhaps liver, but nanobeads localize nonetheless to these organs in whole-body imaging of near-infrared fluorescent (NIRF) beads (Fig. 2A and fig. S3, A to D) by interactions that are likely similar to those that promote RBC adhesion in spleen and greatly increase splenic hematocrit (*24*). Recombinant mCD47 also enhanced persistence in circulation and again showed moderate but suppressed splenic localization (fig. S3B). For additional insight into persistent circulation and potential therapeutic application of Self-nanoparticles, human-derived A549 lung adenocarcinoma epithelial cells were grafted into the flanks of NSG mice, NIRF nanobeads were injected into the tail veins weeks later, and the tumors were imaged both in vivo and ex vivo to quantify accumulated signal. As early as 10 min post injection, hCD47 and Self nanobeads gave mean tumor intensities twice those of noninjected mice, whereas control beads gave background-level signal (Fig. 2B). With hCD47 and Self nanobeads, the fluorescence at every time point is statistically similar but significantly higher than that of control beads ($P < 0.05$), and the increase fits first-order kinetics ($\tau = 52$ min), consistent with enhanced perfusion and progressive clearance. At 40 min, both hCD47 and Self nanobeads give higher signals than controls, ~10 to 20 times as much, and a second injection of hCD47 nanobeads after 2 hours showed a similar signal increase (fig. S4, A and B). Tumor accumulation fits a first-order process (with $\tau = 52$ min giving $T = \tau \ln 2 = 36$ min) that is much faster than control beads ($T = 210$ min), and both time scales are similar to those obtained for persistent circulation (Fig. 1C), consistent with the hypothesis that enhanced tumor signal results from persistent circulation.

After in situ imaging, tumors and other major organs were subsequently excised and imaged ex vivo (fig. S4C). The Self beads and hCD47 beads show at least 16- to 22-fold enhancement of the very low signals obtained with control beads either with or without Scrambled peptide, with no statistical difference between Self peptide and hCD47. The fraction of nanobeads in the blood within the tumor is small (fig. S4C, inset), and so the majority of signal derives from beads that have accumulated in the tumor, most likely by enhanced permeation and retention (EPR) through the leaky vasculature that is characteristic of many solid tumors (*25*). On the basis of these results, the hydrophobic anticancer drug paclitaxel (Tax) was loaded into the Self nanobeads, as well as into beads with PEG and/or antibody against hCD47 (fig. S5, A and B). The latter antibodies have been used therapeutically to mask self on cancer cells (*26*), but targeting antibodies are a double-edged sword when attached to beads, because they also promote clearance (fig. S1D). Tax-loaded beads that displayed either recombinant hCD47 (fig. S5C) or Self-peptide plus PEG and hCD47-targeting antibody (fig. S5D) consistently shrank tumors more than beads lacking Self. The Self beads also did as well or better than the standard paclitaxel

nanocarrier Cremophore, which is known for its toxic side effects (e.g., fig. S5, E and F). The NSG mouse results thus reveal active suppression of clearance by both hCD47 and Self peptide, which could enhance both tumor imaging and drug delivery.

Flow cytometry enables detailed analysis of the surface of nanobeads sampled from circulation. Although streptavidin-specific IgG remains stably bound, biotinylated hCD47 is partially lost (30% in Fig. 3A, inset bar graph; fig. S6A). Nonetheless, the percentage clearance of nanobeads at 35 min versus the measured density of hCD47 at 35 min fits an inhibition model (Fig. 3A) with the inhibition constant $K_{i,in\text{-}vivo}$ of 110 molecules per 160-nm nanobead. This appears independent of circulating bead number over at least a ~10-fold range (fig. S6B). This $K_{i,in\text{-}vivo}$ corresponds to a density of hCD47 that is 10 times that of the lowest densities reported for hRBCs [~25 hCD47 molecules/μm^2 (*12*)] but ~1/100th of nanobead saturation (e.g., fig. S2A). Although binding of soluble hCD47 to NSG-SIRPα$^+$ phagocytes yields a weak affinity (K_d = 4 μM) (Fig. 3B), lymphocytes, which do not express SIRPα (*27*), show zero binding. In vivo evidence of interaction specificity was also obtained by preinjecting mSIRPα-specific antibody, which blocks hCD47 binding, followed by injection of Self nanobeads; these beads were cleared as if lacking Self (Fig. 3B, inset).

To compare the effective affinity of NSG mouse to human SIRPα (hSIRPα), 10 human polymorphic variants of hSIRPα (*14*) were expressed on Chinese hamster ovary (CHO) cell membranes (*24*). Saturation binding of soluble hCD47 to each variant yielded a 60-fold range of affinities with K_d = 0.08 to 5 μM (Fig. 3C and fig. S7), even though all of the amino acid differences in the variants are outside of the binding interface (*22*). When plotted against the allele frequency of hSIRPα, variants of intermediate affinity (e.g., v1 and v2) are most common, with affinities similar to that for soluble SIRPα(v1) binding to CHO-displayed hCD47 (*23*). hCD47's affinity for NSG-SIRPα phagocytes is within the range of reported hSIRPα variants and so is the affinity of Self peptide for the most common SIRPα variant. Synthesis and simulations of additional peptides (Fig. 3D and fig. S8) reveal a sensitivity of binding to conformation, as well as sequence, and show that a lack of affinity for hSIRPα is predictive of a failure to inhibit in vitro phagocytosis.

Whether phagocytosis of nanoparticles—including viruses—involves mechanisms similar to those for larger particles remains an open question (*28*). Tissue sections show that nanobeads colocalize with macrophages (Fig. 4A, *i*). In cultures of human-derived monocytic cell line THP1 macrophages, nanoparticles are not dense enough to settle and contact cells, but opsonized nanobeads that are added at the same total surface area as microbeads (fig. S9A) are taken up with equal efficiency (Fig. 4A, *ii*, *a*). Myosin-II accumulates at the phagocytic synapse formed with opsonized beads except when hCD47 is attached (Fig. 4A, *ii*, *a* and *b*). Microparticles and microbes (i.e., bacteria) give similar images (fig. S9B), and inhibition of phagocytic uptake by hCD47 is indeed independent of particle size from at least 100 nm to 10 μm (Fig. 4A, *ii*, *c*; and fig. S9C). Similar inhibition of nanobead uptake was found with biotinylated Self peptides, whereas both Scrambled peptide and disulfide-bridged peptide (Self-SS) showed no significant inhibition of uptake (Fig. 3D, *i* and *ii*). The potency of hCD47 is remarkable with $K_{i,in\text{-}vitro} \approx 1.0 \pm 0.3$ molecule per 45,000 nm^2. Equivalently, a nanoparticle of 60-nm radius requires only one CD47 molecule to inhibit uptake. $K_{i,in\text{-}vitro}$ is the same as the lowest densities reported for hRBCs (*12*) and is far smaller than PEG densities needed to enhance nanoparticle circulation through delayed opsonization (e.g., fig. S2A: >1 PEG per 20 nm^2).

IgG-driven uptake is linear in myosin-II activity (*16*). With the nanobeads described here, both drug inhibition of myosin-II and hCD47 inhibit uptake by up to 80% (Fig. 4A, *ii*, *c*). When CD47 binds SIRPα, SIRPα's cytoplasmic tail is hyperphosphorylated to activate SHP1 phosphatase (*29*), which dephosphorylates multiple proteins, including myosin-II (*16*). Inhibition of SHP1 produces the expected increase in phagocytosis of hCD47 beads (Fig. 4A, *ii*, *c*; and fig. S9D), and binding of hCD47 and Self peptide indeed increases phospho-SIRPα (Fig. 4B). Consistent with a common mechanism in vitro and in vivo, uptake of the various nanoparticles by THP1 cells correlates inversely with persistence in NSG mice (Fig. 4C).

Phagocytes pervade all tissues and disease sites, with key roles in recognition and clearance, as well as contributions to pro- and anti-inflammatory responses with cytokine release and oxidative burst. Whether synthetic Self peptides, CD47, or its homologs are displayed on particles, viruses (*30*), or artificial surfaces (*31*), "active stealth" signaling across length scales (fig. S10) offers additional opportunities in application as well as understanding. In particular, the SIRPα polymorphism results suggest that an intermediate affinity for Self is optimal (Fig. 3C) as a trade-off between adhesion that is not too strong ("must let go") and signaling that is not too weak ("don't eat me"). Additional homeostatic self factors seem likely and might similarly be used to further avoid phagocytes and thereby enhance delivery of therapeutics and imaging agents.

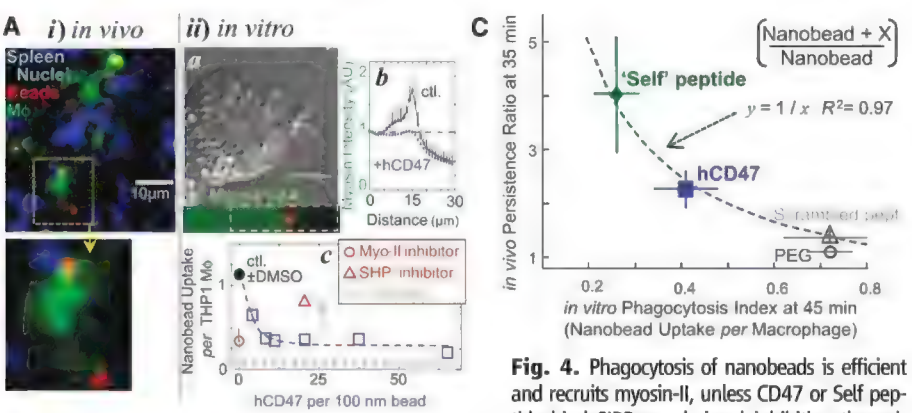


Fig. 4. Phagocytosis of nanobeads is efficient and recruits myosin-II, unless CD47 or Self peptide bind SIRPα and signal inhibition through SHP1. **(A)** Nanobead uptake in NSG mice and in vitro with human-derived THP1 macrophages. (*i*) Splenic macrophages colocalize with nanoparticles in situ. Spleens harvested after 35 to 40 min were frozen, sectioned, fixed, and permeabilized for immunostaining green for macrophages (Mϕ) and red with a secondary antibody against streptavidin-coated opsonized beads [goat antirabbit F(ab')$_2$]. Nuclei are stained blue with Hoechst dye. (*ii*) Phagocytosis of fluorescent 100-nm beads (red) by THP1 cells in vitro was assessed at 45 min by immunostaining cultures that were fixed (but not cell permeabilized) for noningested beads by using secondary antibody against antistreptavidin. Nonmuscle myosin-IIA (*a*, bottom) enriches near the nanoparticle unless hCD47 is on the bead (*b*, plot). Enrichment extends deeply into the cytoplasm (~5 μm) relative to bead size, suggestive of a diffuse signal that directs cytoskeletal assembly. Nanobeads with antistreptavidin are readily engulfed at about 1 bead per cell (*c*), but uptake is inhibited by hCD47 and by inhibition of myosin-IIA with blebbistatin (50 μM). Inhibition of SHP1, downstream of SIRPα, with NSC-87877 blocks the inhibition of uptake by hCD47. Dimethyl sulfoxide (DMSO) is the solvent for the drugs. **(B)** Phosphorylation of hSIRPα tyrosines in THP1 cells upon contact with opsonized nanobeads bearing hCD47 and Self peptide. hSIRPα was immunoprecipitated from cell lysates, and phosphotyrosine was immunoblotted for quantification (n = 3; *P < 0.05). **(C)** Inverse correlation between in vivo persistence ratio at 35 min and in vitro inhibition of phagocytosis by hCD47 and Self peptide at 45 min for 160-nm beads. All data are means ± SEM.

References and Notes

1. D. W. Bartlett, H. Su, I. J. Hildebrandt, W. A. Weber, M. E. Davis, *Proc. Natl. Acad. Sci. U.S.A.* **104**, 15549 (2007).
2. A. L. Klibanov, K. Maruyama, A. M. Beckerleg, V. P. Torchilin, L. Huang, *Biochim. Biophys. Acta* **1062**, 142 (1991).
3. P. J. Photos, L. Bacakova, B. Discher, F. S. Bates, D. E. Discher, *J. Control. Release* **90**, 323 (2003).
4. R. L. Hong *et al.*, *Clin. Cancer Res.* **5**, 3645 (1999).

5. R. Rossin, S. Muro, M. J. Welch, V. R. Muzykantov, D. P. Schuster, *J. Nucl. Med.* **49**, 103 (2008).
6. J. K. Armstrong *et al.*, *Cancer* **110**, 103 (2007).
7. M. J. Turk, D. J. Waters, P. S. Low, *Cancer Lett.* **213**, 165 (2004).
8. P. A. Oldenborg *et al.*, *Science* **288**, 2051 (2000).
9. A. A. Bentley, J. C. Adams, *Mol. Biol. Evol.* **27**, 2187 (2010).
10. E. J. Brown, W. A. Frazier, *Trends Cell Biol.* **11**, 130 (2001).
11. L. J. Bruce *et al.*, *Blood* **101**, 4180 (2003).
12. I. Mouro-Chanteloup *et al.*, *Blood* **101**, 338 (2003).
13. S. Subramanian, R. Parthasarathy, S. Sen, E. T. Boder, D. E. Discher, *Blood* **107**, 2548 (2006).
14. K. Takenaka *et al.*, *Nat. Immunol.* **8**, 1313 (2007).
15. T. Strowig *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **108**, 13218 (2011).
16. R. K. Tsai, D. E. Discher, *J. Cell Biol.* **180**, 989 (2008).
17. F. Turrini, F. Mannu, P. Arese, J. Yuan, P. S. Low, *Blood* **81**, 3146 (1993).
18. D. Wilflingseder *et al.*, *J. Immunol.* **178**, 7840 (2007).
19. M. Lundqvist *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **105**, 14265 (2008).
20. D. Cox, S. Greenberg, *Semin. Immunol.* **13**, 339 (2001).
21. A. M. Glodek *et al.*, *Blood* **116**, 6063 (2010).
22. D. Hatherley *et al.*, *Mol. Cell* **31**, 266 (2008).
23. S. Subramanian, E. T. Boder, D. E. Discher, *J. Biol. Chem.* **282**, 1805 (2007).
24. I. C. MacDonald, E. E. Schmidt, A. C. Groom, *Microvasc. Res.* **42**, 60 (1991).
25. Y. Matsumura, H. Maeda, *Cancer Res.* **46**, 6387 (1986).
26. S. B. Willingham *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **109**, 6662 (2012).
27. M. Seiffert *et al.*, *Blood* **94**, 3633 (1999).
28. J. A. Swanson, A. D. Hoppe, *J. Leukoc. Biol.* **76**, 1093 (2004).
29. T. Matozaki, Y. Murata, H. Okazawa, H. Ohnishi, *Trends Cell Biol.* **19**, 72 (2009).
30. C. M. Cameron, J. W. Barrett, M. Mann, A. Lucas, G. McFadden, *Virology* **337**, 55 (2005).
31. S. J. Stachelek *et al.*, *Biomaterials* **32**, 4317 (2011).

Acknowledgments: The assistance of A. Secreto, J. Glover, and G. Danet-Desnoyers of the Stem Cell Xenograft Core at the University of Pennsylvania is very gratefully acknowledged as is technical assistance of K. Hsu, P. Bhoorasingh, and V. Carnevale. Support from the NIH (R01-EB007049, R01-HL062352, P01-DK032094, NCATS-8UL1TR000003, P30-DK090969) and NSF (Materials Research Science and Engineering Center, and Nano Science and Engineering Center-Nano Bio Interface Center) is also very gratefully acknowledged. P.L.R., D.A.P., and D.E.D. are authors on a patent applied for by the University of Pennsylvania on the Self peptide with first disclosure filed on 16 August 2010.

Supplementary Materials
www.sciencemag.org/cgi/content/full/339/6122/971/DC1
Materials and Methods
Figs. S1 to S10
Table S1
References (*32*–*44*)

31 August 2012; accepted 19 December 2012
10.1126/science.1229568

Caspase-11 Protects Against Bacteria That Escape the Vacuole

Youssef Aachoui,[1,2]† Irina A. Leaf,[3]† Jon A. Hagar,[1,2]† Mary F. Fontana,[4]* Cristine G. Campos,[1] Daniel E. Zak,[3] Michael H. Tan,[4] Peggy A. Cotter,[1] Russell E. Vance,[4] Alan Aderem,[3] Edward A. Miao[1,2]‡

Caspases are either apoptotic or inflammatory. Among inflammatory caspases, caspase-1 and -11 trigger pyroptosis, a form of programmed cell death. Whereas both can be detrimental in inflammatory disease, only caspase-1 has an established protective role during infection. Here, we report that caspase-11 is required for innate immunity to cytosolic, but not vacuolar, bacteria. Although *Salmonella typhimurium* and *Legionella pneumophila* normally reside in the vacuole, specific mutants (*sifA* and *sdhA*, respectively) aberrantly enter the cytosol. These mutants triggered caspase-11, which enhanced clearance of *S. typhimurium sifA* in vivo. This response did not require NLRP3, NLRC4, or ASC inflammasome pathways. *Burkholderia* species that naturally invade the cytosol also triggered caspase-11, which protected mice from lethal challenge with *B. thailandensis* and *B. pseudomallei*. Thus, caspase-11 is critical for surviving exposure to ubiquitous environmental pathogens.

Canonical inflammasomes, such as NLRP3, NLRC4, and AIM2, are cytosolic sensors that detect pathogens or danger signals and activate caspase-1, which leads to secretion of the proinflammatory cytokines interleukin-1β (IL-1β) and IL-18 and pyroptosis, a form of programmed cell death (*1*). Pyrin domain–containing inflammasomes, including NLRP3, signal through the ASC adaptor protein to recruit caspase-1 (fig. S1). Many diverse agonists cause cytosolic perturbations that are detected through NLRP3; however, the underlying mechanisms remain obscure (*2*). In contrast, the CARD domain–containing inflammasome NLRC4 can signal directly to caspase-1, which results in pyroptosis, as well as indirectly through ASC to promote IL-1β and IL-18 secretion (fig. S1) (*1*, *3*). NLRC4 detects bacterial flagellin and type III secretion system (T3SS) rod or needle components within the macrophage cytosol (*4*–*6*). Together, NLRC4- and the ASC-dependent inflammasomes account for all known canonical caspase-1 activation pathways.

Burkholderia pseudomallei is a Gram-negative bacterium endemic to Southeast Asia that causes melioidosis and is a potential biologic weapon (*7*). *B. pseudomallei* uses a T3SS to escape the phagosome and to replicate in the cytosol. NLRC4 and NLRP3 both detect *B. pseudomallei*, which promotes IL-1β secretion from murine bone marrow–derived macrophages (BMMs) (*8*) (Fig. 1A). Despite encoding many of the same virulence factors as *B. pseudomallei*, including T3SS and T6SS, the closely related *B. thailandensis* is far less virulent (*9*). We therefore hypothesized that NLPR3 and NLRC4 also detect *B. thailandensis*, and indeed, NLRP3 and NLRC4 accounted for all IL-1β secretion in response to *B. thailandensis* (Fig. 1B). We next determined whether inflammasome activation is critical to survival after *B. thailandensis* challenge using caspase-1–deficient mice. Kayagaki *et al.* recently showed that all existing caspase-1–deficient mice also lack caspase-11 because of the backcrossing of a mutant *Casp11* allele from 129 into C57BL/6 mice (*10*). Inflammasome detection was critical for resistance to *B. thailandensis*, as $Casp1^{-/-}Casp11^{-/-}$ animals succumbed to the infection (Fig. 1C and fig. S2A). In contrast, wild-type C57BL/6 mice survived high-dose intraperitoneal (i.p.) or intranasal (i.n.) challenge (Fig. 1C and fig. S2A). To our surprise, $Nlrc4^{-/-}Asc^{-/-}$ mice that are deficient in all known canonical inflammasomes were also resistant (Fig. 1D and fig. S2B). This indicated that an unknown signaling pathway provides protection via either caspase-1 or -11 (see pathway schematic in fig. S1). Resistance to *B. thailandensis* was at least partially independent of IL-1β and IL-18, depending on the route of infection (Fig. 1E and fig. S2C), which suggested that both cytokines and pyroptosis can contribute to protection. We therefore examined pyroptosis in vitro and found that cytotoxicity in response to *B. thailandensis* was impaired in $Casp1^{-/-}Casp11^{-/-}$ BMMs (Fig. 1F). Consistent with our in vivo data, pyroptosis in vitro did not require *Nlrc4* or *Asc* (Fig. 1F). *B. pseudomallei* similarly triggered pyroptosis in $Nlrc4^{-/-}Asc^{-/-}$ macrophages (Fig. 1G). These results indicate that a pyroptosis-inducing pathway distinct from all known canonical inflammasomes detects *B. thailandensis* and protects against lethal infection.

Inflammasomes discriminate pathogens from nonpathogens by detecting contamination or perturbation of the cytosolic compartment (*11*). The *B. thailandensis* T3SS facilitates bacterial access to the cytosol and was required for induction of pyroptosis, whereas the virulence-associated T6SS was dispensable (Fig. 2A). We therefore hypothesized that macrophages detect vacuolar lysis or release of bacteria into the cytosol.

In order to establish their intracellular vacuolar growth niche, *Salmonella typhimurium* and *Legionella pneumophila* use T3SS and T4SS, respectively, to translocate effector proteins that work in concert to maintain the stability of these altered, bacteria-containing vacuoles (*12*–*14*). Loss of the *S. typhimurium* SifA or *L. pneumophila* SdhA effectors causes rupture of the vacuole and release of bacteria into the cytosol (*15*–*17*). *S. typhimurium* uses two distinct T3SSs encoded by the *Salmonella* pathogenicity island 1 (SPI1) and SPI2; these two T3SSs translocate distinct batteries

[1]Department of Microbiology and Immunology, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA. [2]Lineberger Comprehensive Cancer Center, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA. [3]Seattle Biomedical Research Institute, Seattle, WA 98109, USA. [4]Department of Molecular and Cell Biology, Division of Immunology and Pathogenesis, University of California, Berkeley, Berkeley, CA 94720, USA.

*Present address: Division of Experimental Medicine, University of California, San Francisco, San Francisco, CA 94122, USA.
†These authors contributed equally to this manuscript.
‡To whom correspondence should be addressed. E-mail: emiao@med.unc.edu

of effectors, such as SifA by SPI2 (*18*). Although *S. typhimurium*–expressing SPI1 and flagellin are readily detected by NLRC4 (*19*, *20*), bacteria grown under conditions that mimic the vacuolar environment express SPI2 and repress flagellin, which minimizes canonical inflammasome detection (*1*, *11*, *21*). Infection of BMMs with *S. typhimurium* that lacked *sifA*, however, significantly increased IL-1β secretion and pyroptosis (Fig. 2, B and C). IL-1β secretion was dependent on canonical inflammasomes (Fig. 2B), whereas pyroptosis was still observed in $Nlrc4^{-/-}Asc^{-/-}$ and $Nlrp3^{-/-}Nlrc4^{-/-}$ macrophages (Fig. 2C). Furthermore, the NLRC4 inflammasome agonist flagellin was not required for these responses (Fig. 2, D and E). Thus, macrophages detect *S. typhimurium* when it aberrantly enters the cytosol and activates pyroptosis independently of all known canonical inflammasomes.

L. pneumophila also translocates flagellin through its T4SS. Thus, *L. pneumophila* mutants lacking flagellin (Δ*flaA*) evaded NLRC4 detection (Fig. 2F) (*2*). In contrast, *L. pneumophila* Δ*flaA* Δ*sdhA* mutants induce caspase-1 activation (*16*, *17*), IL-1β secretion (*17*), and pyroptosis (Fig. 2F) (*17*). The AIM2-ASC canonical inflammasome has been implicated in *L. pneumophila* Δ*flaA* Δ*sdhA*–induced IL-1β secretion, likely by detecting DNA released from bacteria lysing in the cytosol; however, the role of AIM2-ASC in pyroptosis was not examined (*17*). Analogous to *S. typhimurium* Δ*sifA*, *L. pneumophila* Δ*flaA* Δ*sdhA* induced pyroptosis in the absence of flagellin and ASC (Fig. 2G), which rules out all canonical inflammasomes in triggering pyroptosis under these infection conditions. These data demonstrate that diverse bacteria are detected in the cytosol.

Because IL-1β secretion required the canonical inflammasomes, whereas pyroptosis did not, we hypothesized that cell death is triggered by a distinct mechanism mediated by caspase-11. Like caspase-1, caspase-11 is an inflammatory caspase that can directly trigger pyroptosis (fig. S1). Caspase-11 can also promote IL-1β secretion dependent upon NLRP3, ASC, and caspase-1 (*10*, *22*–*24*). Because caspase-1 is activated by recruitment to an oligomerized platform known as the inflammasome, Kayagaki *et al.* hypothesized that a similar oligomeric structure would activate caspase-11, which they termed the noncanonical inflammasome (*10*). Although the cholera toxin B subunit and many different Gram-negative bacteria can trigger caspase-11 activation in vitro (*10*, *22*–*24*), the nature of the physiologic stimulus that activates caspase-11 during infection remains uncertain.

Caspase-11 activation requires priming through a Toll-like receptor 4 (TLR4)–TRIF–STAT1 pathway (*10*, *22*–*24*) (TRIF is a TLR adaptor protein; STAT1 is an interferon receptor signaling protein). Consistent with this, $Tlr4^{-/-}$ and $Trif^{-/-}$ macrophages did not undergo pyroptosis after *S. typhimurium* Δ*sifA* infection, whereas cell death was observed in macrophages deficient in the other TLR4 adaptor, Myd88 (Fig. 3A). This dependence could be overcome by priming the macrophages with interferon-γ (IFN-γ) (Fig. 3A), which signals through STAT1. IFN-γ or lipopolysaccharide (LPS) priming significantly increased the sensitivity of macrophages to *S. typhimurium* Δ*sifA* (Fig. 3A and fig. S3A). These priming effects correlated with increased caspase-11 expression (fig. S3, B and C) but could also be mediated by enhancing aberrant vacuolar rupture. We used retroviruses to complement $Casp1^{-/-}Casp11^{-/-}$ macrophages with either *Casp1* or *Casp11* to determine which was involved. Caspase-11 alone promoted pyroptosis without IL-1β secretion after *B. thailandensis* infection, whereas caspase-1 enabled both responses (Fig. 3B). This is consistent with *B. thailandensis* detection through NLRC4- and/or NLRP3-activating caspase-1 (*8*) and an additional pathway activating caspase-11. In contrast, the responses to *S. typhimurium* Δ*sifA or L. pneumophila* Δ*flaA* Δ*sdhA* acted through caspase-11 but not caspase-1 (Fig. 3, C and D). We further confirmed that caspase-11 was responsible for the cell death observed in $Nlrc4^{-/-}Asc^{-/-}$ macrophages, using microRNA-adapted short hairpin RNA (shRNAmir) (Fig. 3, E and F, and fig. S3E). Finally, $Casp11^{-/-}$ BMMs revealed that caspase-11 was required for pyroptosis after *B. thailandensis*, *S. typhimurium* Δ*sifA*, and *L. pneumophila* Δ*flaA* Δ*sdhA* (Fig. 3, G to I). Although a previous report suggested that NLRC4 signals through caspase-11 to alter phagosomal trafficking (*25*), we saw no evidence that NLRC4 contributes to caspase-11–dependent cell death (Figs. 1F and 2D and fig. S4). Pyroptosis initiated by caspase-11 was morphologically similar to pyroptosis triggered by caspase-1 (fig. S5, A and B). Therefore, macrophages activate caspase-11 in response to cytosolic *B. thailandensis*, *S. typhimurium*, or *L. pneumophila* (fig. S1).

S. typhimurium Δ*sifA* is attenuated (*15*), which has previously been attributed to the role of SifA in coordinating intracellular trafficking of the *Salmonella*-containing vacuole. We hypothesized that this attenuation was actually due to innate immune detection though caspase-11. Indeed, *S. typhimurium* Δ*sifA* was mildly attenuated in

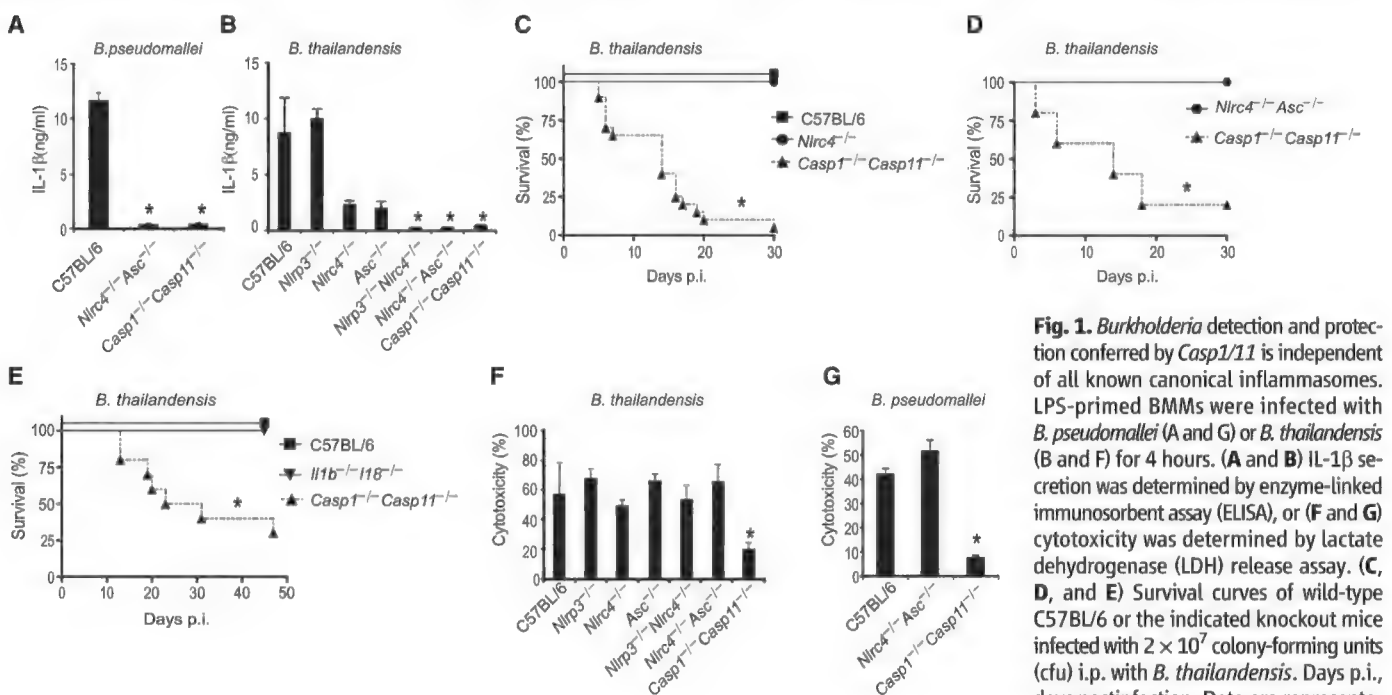


Fig. 1. *Burkholderia* detection and protection conferred by *Casp1/11* is independent of all known canonical inflammasomes. LPS-primed BMMs were infected with *B. pseudomallei* (A and G) or *B. thailandensis* (B and F) for 4 hours. (**A** and **B**) IL-1β secretion was determined by enzyme-linked immunosorbent assay (ELISA), or (**F** and **G**) cytotoxicity was determined by lactate dehydrogenase (LDH) release assay. (**C**, **D**, and **E**) Survival curves of wild-type C57BL/6 or the indicated knockout mice infected with 2×10^7 colony-forming units (cfu) i.p. with *B. thailandensis*. Days p.i., days postinfection. Data are representative of at least three (A), (B), (F), and (G) or of two (D) and (E) experiments. (C) Data are pooled from three experiments. For number of mice in each panel, see table S2. Statistically significant differences with respect to controls are indicated (Student's *t* test or log-rank test for survival; $*P \leq 0.05$).

C57BL/6 mice as expected, but this was not replicated in *Casp1*$^{-/-}$*Casp11*$^{-/-}$ mice (Fig. 4, A and B). We next determined the relative clearance of *S. typhimurium* Δ*sifA* during co-infection with wild-type *S. typhimurium*, a more quantitative measure of virulence than lethal challenge. We recovered 16 times as much wild-type *S. typhimurium* as *S. typhimurium* Δ*sifA* from C57BL/6 mice; however, from *Casp11*$^{-/-}$ mice, we only recovered four times as many wild-type bacteria (Fig. 4C). This indicates that caspase-11 clears *S. typhimurium* Δ*sifA* in vivo; in contrast, wild-type *S. typhimurium* effectively evades caspase-11 (*23*) by remaining within the vacuole. The remaining *S. typhimurium* Δ*sifA* attenuation likely reflects the role of *sifA* as a virulence factor promoting intracellular replication. Moreover, all known canonical inflammasomes were dispensable for *S. typhimurium* Δ*sifA* clearance, as were IL-1β and IL-18 (Fig. 4D), which implicated pyroptosis as the mechanism of clearance. Clearance of bacteria after pyroptosis is mediated by neutrophils through generation of reactive oxygen (*21*). Consistent with this, NADPH oxidase–

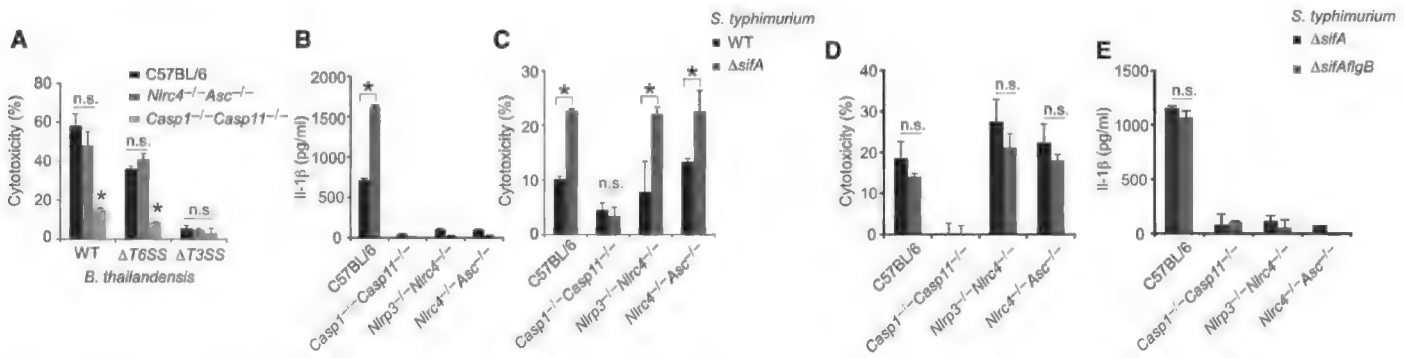
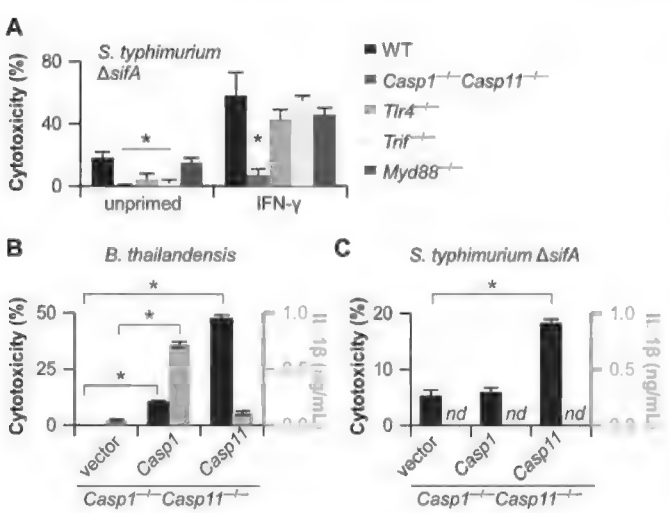


Fig. 2. Diverse cytosolic bacteria activate pyroptosis independent of NLRC4, NLRP3, and ASC. (**A**) LPS-primed BMMs were infected for 4 hours with either *B. thailandensis* or the indicated mutants, and cytotoxicity was determined. BMMs were infected for 8 hours with (**B**) *S. typhimurium* or *S. typhimurium* Δ*sifA* or (**D**) *S. typhimurium* Δ*sifA* or *S. typhimurium* Δ*sifA flgB*, and cytotoxicity was determined. WT, wild type. LPS-primed BMMs were infected for 8 hours with (**C**) *S. typhimurium* or *S. typhimurium* Δ*sifA* or (**E**) *S. typhimurium* Δ*sifA* or *S. typhimurium* Δ*sifA flgB*, and IL-1β secretion was determined. (**F** and **G**) Cytotoxicity in wild-type or *Asc*$^{-/-}$ BMMs infected for 4 hours with *L. pneumophila*, *L. pneumophila* Δ*flaA*, or *L. pneumophila* Δ*sdhA* Δ*flaA*. Cytotoxicity was determined by LDH release and IL-1β secretion by ELISA. Data are representative of at least three experiments. Statistically significant differences with respect to controls are indicated (Student's *t* test; $*P \leq 0.05$; n.s., $P > 0.05$).



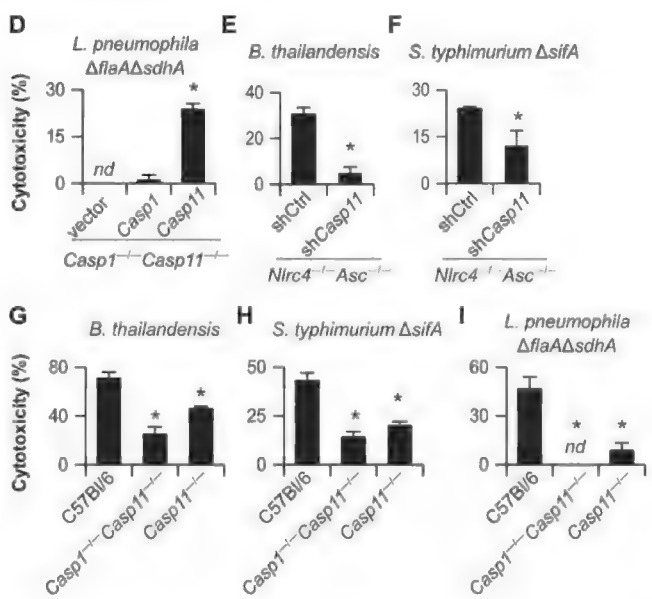


Fig. 3. Caspase-11 mediates pyroptosis after infection by cytosolic bacteria. Macrophage cytotoxicity and IL-1β secretion were determined after infection with *S. typhimurium* Δ*sifA* (8 hours), *L. pneumophila* Δ*flaA* Δ*sdhA* (4 hours), or *B. thailandensis* (4 hours). (**A**) C57BL/6, *Casp1*$^{-/-}$*Casp11*$^{-/-}$, *Tlr4*$^{-/-}$, *Trif*$^{-/-}$, and *Myd88*$^{-/-}$ BMMs infected with *S. typhimurium* Δ*sifA* with or without IFN-γ priming before infection. (**B** and **C**) Retroviral transduction was used to complement *Casp1* or *Casp11* in *Casp1*$^{-/-}$*Casp11*$^{-/-}$ immortalized BMMs. Macrophages were primed with LPS (B) or IFN-γ (C), and responses to *B. thailandensis* (B) or *S. typhimurium* Δ*sifA* (C) infection were examined. (**D**) Control or complemented *Casp1*$^{-/-}$*Casp11*$^{-/-}$ BMMs infected with *L. pneumophila* Δ*flaA* Δ*sdhA*. (**E** and **F**) Retroviral transduction was used to introduce control or *Casp11*-targeting of shRNAmir into *Nlrc4*$^{-/-}$*Asc*$^{-/-}$ immortalized BMMs. Macrophages were primed overnight with LPS (E) or IFN-γ (F) and then infected as indicated. (**G** to **I**) C57BL/6, *Casp1*$^{-/-}$*Casp11*$^{-/-}$, and *Casp11*$^{-/-}$ BMMs infected with *B. thailandensis* (G), *S. typhimurium* Δ*sifA* (H), or *L. pneumophila* Δ*flaA* Δ*sdhA* (I). Data are representative of at least three (A to C), (E), (G), and (H) or of two (D), (F), and (I) independent experiments. Statistically significant differences with respect to controls are indicated (Student's *t* test; $*P \leq 0.05$). *nd*, none detected.

deficient *p47phox*$^{-/-}$ mice were also defective for clearance of *S. typhimurium* Δ*sifA* (Fig. 4D). However, TLR4 and IFN-γ were not required (Fig. 4E), which suggests that there is redundant priming of caspase-11 pathways in vivo. Therefore, caspase-11 protects mice from *S. typhimurium* Δ*sifA*, and because IL-1β and IL-18 are not required, pyroptosis is likely to be the mechanism of bacterial clearance in this case.

We next examined the susceptibility of *Casp11*$^{-/-}$ mice to the naturally cytosolic pathogens *B. thailandensis* and *B. pseudomallei*. Although C57BL/6 mice are resistant to *B. thailandensis* infection, *Casp11*$^{-/-}$ mice succumbed (Fig. 4F). Likewise, *Casp11*$^{-/-}$ mice succumbed to *B. pseudomallei* infection, whereas C57BL/6 mice survived (Fig. 4G). Because *Nlrc4*$^{-/-}$ mice are also susceptible to *B. pseudomallei* infection (*8*), we conclude that both caspase-1 and caspase-11 play critical roles in limiting *B. pseudomallei* infection.

Collectively, these data demonstrate, for the first time, that caspase-11 protects animals from lethal infection by bacteria that have the ability to invade the cytosol. This could be critical for defense against ubiquitous environmental bacteria such as *B. thailandensis* that encode virulence factors but that have not evolved to evade caspase-11 detection. It will be interesting to determine whether caspase-11 is activated in response to the process of vacuolar rupture or the presence of bacteria within the cytosol. Caspase-11 also responds to vacuolar bacteria under delayed kinetics, but such responses have not been shown to provide protection from infection in vivo (*10*, *22*–*24*). LPS-induced septic shock is mediated by caspase-11 (*10*), which suggests that caspase-11 can be activated by other mechanisms besides cytosol-localized bacteria. Thus, we propose that caspase-11 provides protection against pathogens, but is dysregulated during overwhelming infection and so contributes to septic shock and mortality. It will be interesting to determine whether caspase-11 triggers eicosanoid secretion, as is seen for caspase-1, and whether these mediators contribute to septic shock (*26*). The identity of the hypothesized noncanonical inflammasome(s) that activate caspase-11 and the precise nature of the activating signal will shed more light on the mechanisms by which caspase-11 can both promote innate immunity and exacerbate immunopathology. These insights may lead to novel therapies to treat infection and sepsis.

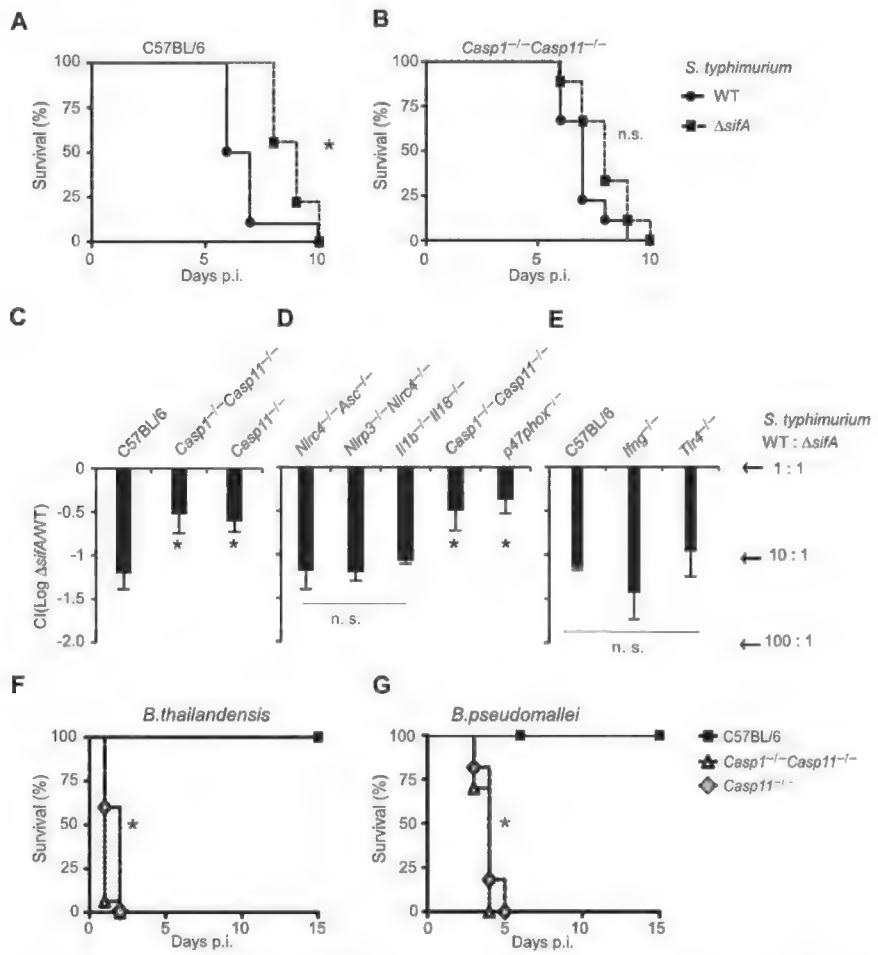


Fig. 4. Caspase-11 protects against cytosolic bacteria in vivo. (**A** and **B**) *S. typhimurium* or *S. typhimurium* Δ*sifA* were injected i.p. into C57BL/6 (1000 cfu) or *Casp1*$^{-/-}$*Casp11*$^{-/-}$ mice (250 cfu), and survival was monitored. (**C** to **E**) The indicated mice were infected with 5×10^4 cfu of both wild-type *S. typhimurium* and *S. typhimurium* Δ*sifA* marked with ampicillin or kanamycin resistance, respectively. Bacterial loads from three to four mice per genotype were determined 48 hours later, and the competitive index was calculated [CI = log(*S. typhimurium* Δ*sifA*/*S. typhimurium*)]. A CI of −1 corresponds to 10 cfu of *S. typhimurium* for every 1 cfu of *S. typhimurium* Δ*sifA*. (**F** and **G**) C57BL/6, *Casp1*$^{-/-}$*Casp11*$^{-/-}$, or *Casp11*$^{-/-}$ mice were infected with (F) 2×10^7 cfu mouse passaged *B. thailandensis* i.p. or (G) 100 cfu *B. pseudomallei* i.n. (A), (B), (F), and (G) Data are pooled from two independent experiments. (C to E) Representative of two experiments. For the number of mice in each panel, see table S2. Statistically significant differences with respect to controls are indicated (Student's *t* test or log-rank test for survival; $^*P \leq 0.05$; n.s., $P > 0.05$).

References and Notes

1. E. A. Miao, J. V. Rajan, A. Aderem, *Immunol. Rev.* **243**, 206 (2011).
2. L. Franchi, R. Muñoz-Planillo, G. Núñez, *Nat. Immunol.* **13**, 325 (2012).
3. P. Broz, J. von Moltke, J. W. Jones, R. E. Vance, D. M. Monack, *Cell Host Microbe* **8**, 471 (2010).
4. E. A. Miao *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **107**, 3076 (2010).
5. Y. Zhao *et al.*, *Nature* **477**, 596 (2011).
6. E. M. Kofoed, R. E. Vance, *Nature* **477**, 592 (2011).
7. W. J. Wiersinga, B. J. Currie, S. J. Peacock, *N. Engl. J. Med.* **367**, 1035 (2012).
8. I. Ceballos-Olvera, M. Sahoo, M. A. Miller, L. Del Barrio, F. Re, *PLoS Pathog.* **7**, e1002452 (2011).
9. W. J. Wiersinga, T. van der Poll, N. J. White, N. P. Day, S. J. Peacock, *Nat. Rev. Microbiol.* **4**, 272 (2006).
10. N. Kayagaki *et al.*, *Nature* **479**, 117 (2011).
11. E. A. Miao, J. V. Rajan, *Front. Microbiol.* **2**, 85 (2011).
12. N. Schroeder, L. J. Mota, S. Méresse, *Trends Microbiol.* **19**, 268 (2011).
13. S. Shoma *et al.*, *Infect. Immun.* **76**, 1547 (2008).
14. J. Ge, F. Shao, *Cell. Microbiol.* **13**, 1870 (2011).
15. C. R. Beuzón *et al.*, *EMBO J.* **19**, 3235 (2000).
16. E. A. Creasey, R. R. Isberg, *Proc. Natl. Acad. Sci. U.S.A.* **109**, 3481 (2012).
17. J. Ge, Y.-N. Gong, Y. Xu, F. Shao, *Proc. Natl. Acad. Sci. U.S.A.* **109**, 6193 (2012).
18. J. van der Heijden, B. B. Finlay, *Future Microbiol.* **7**, 685 (2012).
19. L. Franchi *et al.*, *Nat. Immunol.* **7**, 576 (2006).
20. E. A. Miao *et al.*, *Nat. Immunol.* **7**, 569 (2006).
21. E. A. Miao *et al.*, *Nat. Immunol.* **11**, 1136 (2010).
22. P. Gurung *et al.*, *J. Biol. Chem.* **287**, 34474 (2012).
23. P. Broz *et al.*, *Nature* **490**, 288 (2012).
24. V. A. K. Rathinam *et al.*, *Cell* **150**, 606 (2012).
25. A. Akhter *et al.*, *Immunity* **37**, 35 (2012).
26. J. von Moltke *et al.*, *Nature* **490**, 107 (2012).

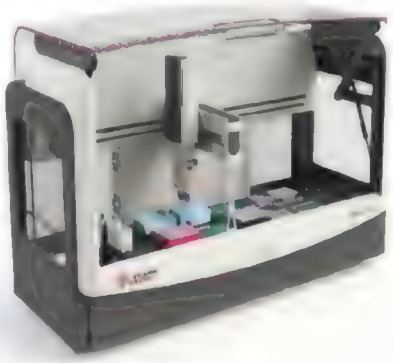
Acknowledgments: The authors thank V. Dixit for sharing mice (a material transfer agreement is on file) and M. Heise for sharing mice and S. Miller, J. Mougous, and H. Schweizer for sharing bacterial strains. We also thank D. Rodriguez and Z. Zhou for managing mouse colonies. The data presented in this manuscript are tabulated in the main paper and in the supplementary materials. This work was supported by NIH grants AI097518 (E.A.M.), AI057141 (E.A.M. and A.A.), AI065359 (P.A.C.), AI075039 (R.E.V.), AI080749 (R.E.V.), and AI063302 (R.E.V.); Investigator Awards from the Burroughs Wellcome Fund and Cancer Research Institute (R.E.V.); and an NSF graduate fellowship (M.F.F.).

Supplementary Materials

www.sciencemag.org/cgi/content/full/science.1230751/DC1
Materials and Methods
Figs. S1 to S5
Tables S1 and S2
References (*27*–*48*)

26 September 2012; accepted 20 December 2012
Published online 24 January 2013;
10.1126/science.1230751

LIQUID HANDLING AUTOMATION



Featuring intelligent liquid handling that adapts as processes require, the Biomek 4000 Laboratory Automation Workstation incorporates easy-to-use, icon-driven software and an enhanced work surface with interchangeable tools. Modular deck configuration allows the work surface to be set up with eight standard positions, and is expandable to 12 positions. New single and multichannel 1,000 mL pipetting tools offer higher throughput for assays using volumes greater than 200 mL and up to 1,000 mL. Method building is simple, and powerful editing features allow specialized and one-off applications to be readily addressed. Most pipetting functions are available through transfer and combine steps, while the flexibility to fine-tune individual pipetting steps is retained. Pipetting templates and liquid-type editors are standard in the software and make adapting pipetting for any liquid straightforward. An intuitive software interface provides icons for liquid handling, labware movement, and control of external devices.

Beckman Coulter

For info: 800-742-2345 | www.biomek4k.com

AUTOMATED NUCLEIC ACID EXTRACTION WORKSTATION

The new Automated Nucleic Acid Extraction WorkStation offers consistent, high-speed purifications by combining several Thermo Fisher tools into one turnkey solution. Designed to provide cost- and time-effective nucleic acid extraction, users can process samples from virtually any source, while ensuring accuracy and sample integrity. The system is ideal for a variety of downstream applications including sequencing, cloning, and both polymerase chain reaction (PCR) and qPCR, and allows instruments to be used in a complete workflow, alone, or even interchanged while the system is running. Controlled and monitored by the Thermo Scientific Momentum scheduling software, the Automated Nucleic Acid Extraction Workstation and other networked tools can be easily programmed to ensure the simultaneous running of multiple processes. When leaving the instruments unattended, the software enhances user confidence by providing accelerated simulation to test required processes in advance, isolating problems, and delivering recovery options for the workflow.

Thermo Fisher Scientific

For info: 905-332-2000 | www.thermoscientific.com

LIQUID HANDLING SYSTEM

The Preddator System Liquid Handling System, which features unique dispense head technology, achieves less than 1% CVs at nanoliter volumes, ensuring scientists can save time and reagent costs by generating consistent low-volume microplate assay results time after time. The Preddator is the only automated commercial nanoliter microplate dispenser that can reliably pipette difficult to dispense substances such as DMSO, oils, detergents, and gels as well as cells, proteins, and other biological samples. Preddator's unique dispense head is designed with a solenoid to work continuously, saving scientists time as they can leave Preddator unattended when dispensing viscous reagents or cells, without worrying about blockages and inaccurate assay data which can result from inconsistent dispensing.

Redd&Whyte

For info: +44-(0)-1284-703168 | www.reddandwhyte.com

COLONY GROWTH AND PICKING SYSTEM

The fully automated RapidPick Automated Colony Isolation System (ACIS) reduces the total time required to go from a plated bacteria to inoculated wells to bacterial cultures. The RapidPick ACIS brings together multiple instruments in a cost-effective and scalable manner. As a result, several time-consuming steps have been automated that are typically performed manually or offline due to cost, complexity, and other considerations. The RapidPick ACIS allows the user to create a streamlined process for all the instruments to work together, reducing the time required for the process. A unique feature of the RapidPick ACIS is the inclusion of an integrated Micro10X dispenser with auto-prime and auto-rinse features. The Micro10x will automatically fill the plates with fresh media on the fly immediately before inoculation, saving the user valuable time. High throughput runs are possible because the attached PlateCrane EX robotic arm with stacks is able to store and manage multiple microplates in a single run.

Hudson Robotics

For info: 973-376-7400 | www.hudsonrobotics.com

HIGH-CONTENT IMAGING

Replacing the acumen eX3, the new acumen hci (high-content imaging) system is the fastest imager on the market for multiplex assays in up to 1,536-well plates, and the new open-source image format means users have complete flexibility to analyze images on their existing analysis software if required. Together with whole-well imaging, these features make the acumen hci ideal for identifying "hit" wells in cell-based screens. Achieving unparalleled speeds during high throughput screening, acumen hci can image and analyze whole wells for 96- to 1,536-well plates in just eight minutes, at a resolution of 0.5 µm/pixel. This enables the effective screening of high-density plates and facilitates assay miniaturization, increasing throughput and reducing costs. Data is robust and insightful even during rapid screening, and up to three lasers allow for easy multiplexing using a wide range of fluorescent dyes. The new 561 nm laser option offers compatibility with even more dyes, like mCherry.

TTP Labtech

For info: +44-(0)-1763-262626 | www.ttplabtech.com/acumen

Electronically submit your new product description or product literature information! Go to **www.sciencemag.org/products/newproducts.dtl** for more information.

Newly offered instrumentation, apparatus, and laboratory materials of interest to researchers in all disciplines in academic, industrial, and governmental organizations are featured in this space. Emphasis is given to purpose, chief characteristics, and availability of products and materials. Endorsement by *Science* or AAAS of any products or materials mentioned is not implied. Additional information may be obtained from the manufacturer or supplier.

Science Careers Advertising

For full advertising details, go to ScienceCareers.org and click For Employers, or call one of our representatives.

Tracy Holmes
Worldwide Associate Director
Science Careers
Phone: +44 (0) 1223 326525

THE AMERICAS

E-mail: advertise@sciencecareers.org
Fax: 202-289-6742

Tina Burks
East Coast/West Coast/South America
Phone: 202-326-6577

Allyson Rosen
Midwest/Canada/Corporate
Phone: 202-326-6578

Marci Gallun
Sales Administrator
Phone: 202-326-6582

Online Job Posting Questions
Phone: 202-312-6375

EUROPE / INDIA / AUSTRALIA / NEW ZEALAND / REST OF WORLD

E-mail: ads@science-int.co.uk
Fax: +44 (0) 1223 326532

Lucy Nelson
Phone: +44 (0)1223 326527

Kelly Grace
Phone: +44 (0) 1223 326528

JAPAN

Yuri Kobayashi
Phone: +81-(0)90-9110-1719
E-mail: ykobayas@aaas.org

CHINA / KOREA / SINGAPORE / TAIWAN / THAILAND

Ruolei Wu
Phone: +86-1367-1015-294
E-mail: rwu@aaas.org

All ads submitted for publication must comply with applicable U.S. and non-U.S. laws. *Science* reserves the right to refuse any advertisement at its sole discretion for any reason, including without limitation for offensive language or inappropriate content, and all advertising is subject to publisher approval. *Science* encourages our readers to alert us to any ads that they feel may be discriminatory or offensive.

Science Careers
From the journal *Science* AAAS

ScienceCareers.org

Max Planck Institute of Immunobiology and Epigenetics
Max-Planck-Institut für Immunbiologie und Epigenetik

We are looking for a

Postdoctoral Position in Bioinformatics

The Max Planck Institute of Immunobiology and Epigenetics in Freiburg, Germany is offering a Postdoctoral Position in Bioinformatics in the Laboratory of Chromatin Regulation (Head: Dr. Asifa Akhtar). The position is available for an initial two-year appointment with the possibility of extension.

The MPI in Freiburg is an international research institute at the cross-road of Southern Germany, Switzerland and France. The working language is English. State-of-the-art infrastructure and service units, including transgenesis, mass spectrometry, proteomics, flow cytometry, fly and imaging facilities are available.

Your tasks: The research focus of the Akhtar laboratory includes mechanisms underlying chromatin and epigenetic regulation. We are particularly interesed in X chromosomal regulation using flies and mouse models employing multidisciplinary approaches such as genetics, biochemistry, functional genomics as well as cell biology and structural biology.

We are looking for enthusiastic, highly-motivated, science-driven and experienced postdoctoral fellows to join our team to unravel the molecular mechanisms that regulate gene expression.

Your qualifications: Applicants should have a PhD or equivalent doctoral degree with at least 3 years of proven research experience in bioinformatics and analyses of genomewide data (ChIPseq, RNA seq). Prior experience working in Drosophila or mammalian models is highly encouraged. Candidates must have a strong publication record. Furthermore, the ability to work in a team, communication skills and experience in the supervision of graduate students are an asset.

Please submit your application, including a statement of research interests, a CV and three references to our homepage: http://www.ie-freiburg.mpg.de/jobs

We offer: Salaries will be according to postdoctoral fellowships of the Max Planck Society or TVöD and will commensurate with experience.

Application deadline: 31.03.2013

Our institute investigates the molecular basis of the immune response and other topics of the developmental biology, such as the origin and differentiation of the immune cells as well as the development of vertebrate embryo. Another main focus of the institute is Epigenetic. This area deals with inheritable traits, which are not caused by changes in the DNA sequence.

Handicapped applicants with equal qualifications will be given preferential treatment. The Max Planck Society seeks to increase the number of women in areas, where they are underrepresented, and therefore explicitly encourages women to apply. A childcare facility is directly attached to the institute. If you would like to work in a dedicated team, please convince us now by sending us your complete application documents together with your salary expectations and your earliest possible date of joining the institute.

Max Planck Institute of Immunobiology and Epigenetics, Ms Klank

Have we sparked your interest? Please apply online via the jobmarket at our homepage. We are looking forward to getting your complete application documents. http://www.ie-freiburg.mpg.de/jobs

Advancing Science in India

Science in India is on the move in a big way. The government has initiated multibillion dollar investments to kick start research, education, and innovation over the next five years. Though several challenging issues remain for the country, India's best and brightest expats living in the United States and Europe are being enticed back to 'Mother India' with the promise of world-class research infrastructure and solid funding. **By Adarsh Sandhu**

In early 2013, India's government announced an ambitious science, technology, and innovation funding protocol: in the next five years, double its investment in science and technology and, by 2020, drive India's output of scientific publications to be among the top five nations globally. "The government is going to inject $5 billion into science and technology over the next five years," says C.N.R. Rao, the founder of the **Jawaharlal Nehru Centre for Advanced Scientific Research** (JNCASR) and chairman of the Science Advisory Council to the Prime Minister. "This doubles the investment to-date from 1% to 2% of GDP." This increase in funding is aimed at creating jobs, educating technical leaders, and improving the quality of science in this country of 1.2 billion people, he notes.

The announcement is just one of a recent number of nationwide initiatives that have been inaugurated as India seeks to improve its global scientific reputation. The creation of new institutions and universities, opportunities for independent leadership training, and efforts to expand translational research and cultivate a culture of technology transfer are just a few of the federal components encouraging young researchers to set up shop in their homeland. In addition, international alliances, between India and organizations in the United States, United Kingdom, and other countries, are also making an impact in bolstering collaboration across borders and building strong scientific capacity within the subcontinent.

But despite these outreach and funding programs, there are still some challenges that need to be addressed before scientists in India can stand shoulder to shoulder with their counterparts in the West. Recent infrastructure investment programs have successfully produced new facilities and institutions all over India, but this has created a shortage of scientists and experts to run and manage the new universities and research institutes. Specifically, according to university administrators, India needs an estimated 40,000 qualified scientists to fill positions currently vacant. However, there is insufficient talent within India to take up this slack, which is being compounded by current labor laws that can sometimes make hiring foreign nationals a complex and difficult process. (Autonomous institutions have provisions for hiring foreigners, albeit on a non-permanent basis.) Nonetheless, both theoretical and translational research in India is moving forward with areas such as nanotechnology, energy, and health at the forefront.

"Perhaps the most dramatic recent change has been that there is no shortage of funding for research."

Expanding Facilities and Infrastructure

The **Indian Institutes of Technology** (IITs) are among India's most prestigious academic institutions. These autonomous institutes were established in the early 1960s, and the government has since expanded their number from the original five—Kharagpur, Bombay (Mumbai), Madras (Chennai), Kanpur, and Delhi—to a total of 16. This increase reflects the government's new policies to give students from a wider range of social backgrounds the opportunity to study at India's top-tier universities.

Sudhir Chandra, a professor at the Centre for Applied Research in Electronics at IIT Delhi, has witnessed dramatic changes during continued>

Sudhir Chandra

CREDIT: (TOP) ISTOCKPHOTO.COM/ADVENTTR/DEEJPILOT; (BOTTOM) PHOTO BY ADARSH SANDHU

Upcoming Features

Postdocs: Identifying Opportunities—March 8
Regional Focus: Germany—March 22
Cancer Research Careers—March 29

National Centre for Biological Sciences

C.N.R. Rao

"Our faculty search committee [at IIT Delhi] operates all year round, interviewing candidates overseas, and when necessary, offering positions on the spot—this is unprecedented."
—Shiban K. Koul

his 32 years as an academic. "In the early days, the IITs concentrated on education," says Chandra. "Then, a decade or so ago, we started to place a greater emphasis on research. Perhaps the most dramatic recent change has been that there is no shortage of funding for research."

Recent research funding surges have led to the development of high-profile projects including the national nanotechnology network, which includes the approximately $11 million Nanoscale Research Facility (NRF) at IIT Delhi and the $40 million Centre for Nano Science and Engineering (CeNSE) at the prestigious **Indian Institute of Science** (IISc) in Bangalore. The nanotechnology facilities were established to train experts and provide experimental facilities for scientists across all of India.

In fact, India has done much to magnify its infrastructure. According to a January 6, 2013 article in University World News online, in recent years the nation has launched five new Indian institutes of science education and research, eight new IITs, 16 new central universities, 10 new national institutes of technology, six new research and development institutions in biotechnology, and five institutions in other branches including biomimetic materials and solar energy.

"We are building a new 250-faculty campus with an initial thrust on interdisciplinary research in Hyderabad," says Mustansir Barma, theoretical physicist and director of **Tata Institute of Fundamental Research** (TIFR) in Mumbai, one of India's oldest and premier basic research institutes. The new 200-acre campus will be significantly larger than the Mumbai facility. "We have also established the new International Centre for Theoretical Sciences in Bangalore—similar to, but broader in scope than, the Isaac Newton Institute for Mathematical Sciences in Cambridge, which organizes international research programs in pure and applied mathematics." Although, Barma notes, "finding high-quality staff to teach and manage the new centers will be a challenge."

In an effort to fill the abundance of new positions, the government has established programs to court Indians working abroad back to their home country. The Ramalingaswami Re-entry Fellowship, funded by the Department of Biotechnology (DBT) under the Ministry of Science and Technology (MST), is designed to attract highly skilled researchers working overseas in a variety biotechnology disciplines. Fellows can have a Ph.D., M.D., M.Tech, M.VSc, or equivalent degree, and are given a stipend, grants, and even a housing allowance. The equally competitive Ramanujan Fellowships, supported by the Department of Science and Technology (DST), also aims to lure senior scientists and engineers, both originally from India and elsewhere, to the country.

Increasing Career Opportunities

With the increases in funding and rapidly expanding institutions, opportunities are becoming more readily available for scientists who want to work in India. "We currently have about 300 vacancies for faculty," says Shiban K. Koul, deputy director of strategy and planning at IIT Delhi. "Our faculty search committee operates all year round, interviewing candidates overseas, and when necessary, offering positions on the spot—this is unprecedented."

However, the choice of recruits for IITs is limited because India's labor laws do not allow these institutes to hire foreign nationals for tenured positions, and the salaries are not as competitive as those in the United States and European Union.

But not all institutions in India encounter these problems. "IISc is the first destination for returnees to India," says Kamanio Chattopadhyay, professor of materials engineering. "Our commitment to excellence in research, the facilities on campus, and our interdisciplinary approach to education are some of the reasons for this." Indeed, IISc has a unique history as a research institute: it was established in 1909 with the support of the late industrialist Jamsetji Nusserwanji Tata and the Maharaja of Mysore Krishnaraja Wodeyar.

One challenge that is on the minds of both Indian academics and government representatives is the ability to draw and retain talented postdocs and other early-career scholars into India's institutions. Some researchers are less than optimistic about finding a solution to attract postdocs to India when higher remuneration packages are available in the West. "I am finding it very difficult to find well-trained postdocs," confirms Madhusudhan Venkadesan at the **National Centre for Biological Sciences** (NCBS), Bangalore. Although this situation is not uncommon in India, certain national, international, "and institutional programs may be helping to alleviate this problem," notes DNA chemist Yamuna Krishnan at NCBS, who is collaborating with scientists in France and Germany with funding from the United Kingdom's Wellcome Trust.

The Council of Scientific and Industrial Research (CSIR)-Nehru Science Postdoctoral Research Fellowship is one such program designed to engage younger scientists. This postdoc opportunity seeks to identify promising young researchers with innovative ideas and provide them with training to transition into independent research continued>

CREDITS: PHOTOS BY ADARSH SANDHU

Indo-U.S. Science & Technology Forum

IUSSTF

Promoting & catalyzing the Indo-U.S. bilateral collaborations in

• Science • Technology • Engineering • Biomedical Research
• Innovation & Entrepreneurship

Capitalizing on the scientific and technological synergy to build long term partnerships

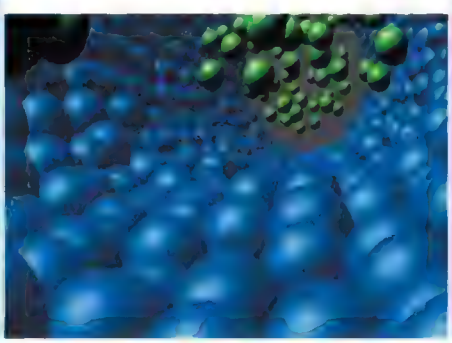
Supporting exciting and enabling science and technology programs

Nurturing contacts between young and mid career scientists through Faculty and Student Fellowships

Conforming to the highest standards of excellence ...

Creating scientific awareness through Bilateral Workshops & Training Programs

Exploring new frontiers of collaboration through Networked Joint Centres



Encouraging public-private partnership through Innovation Promotion Programs

The Indo-U.S. Science and Technology Forum (IUSSTF), established under an agreement between the Governments of India and the United States of America in March 2000, is an autonomous, not for profit society and in India, co-funded and co-governed by both the Governments.

Faculty Fellowship & Student Internship Opportunities

Research Internship in Science and Engineering for American Students

Indo-US Research Fellowship for Indian Researchers

S.N. Bose SCHOLARS PROGRAM

KHORANA Program for SCHOLARS

http://www.indousstf.org

Featured Participants

Indian Institute of Science, Bangalore
www.iisc.ernet.in

Indian Institute of Technology, Delhi
www.iitd.ac.in

Jawaharlal Nehru Centre for Advanced Scientific Research (JNCASR)
www.jncasr.ac.in

National Centre for Biological Sciences (NCBS), Bangalore
www.ncbs.res.in

Tata Institute of Fundamental Research, Mumbai
www.tifr.res.in/index.php/en

University of Wisconsin-Madison (UW) Khorana Program
www.biochem.wisc.edu/faculty/ansari/khorana_program

Additional Resources

Council of Scientific and Industrial Research (CSIR)
rdpp.csir.res.in/csir_acsir

Department of Biotechnology (DBT)
dbtindia.nic.in/index.asp

Department of Science and Technology (DST)
www.dst.gov.in

Indo-U.S. Science and Technology Forum (IUSSTF)
www.indousstf.org

Rajiv Gandhi Centre for Biotechnology
rgcb.res.in

Wellcome Trust/DBT India Alliance
www.wellcomedbt.org

Solar Cells

Both theoretical and translational research in India is moving forward with areas such as nanotechnology, energy, and health at the forefront.

careers. The DBT Rapid Grant for Young Investigators and the DST Swarnajayanti Fellowships Scheme provide comparable support for younger scholars. And the Wellcome Trust/DBT India Alliance, an £80 million initiative funded equally by the Wellcome Trust and the DBT, provides competitive life sciences and biomedical fellowships for postdocs and other early-career scientists. Furthermore, numerous universities and research centers, such as the NCBS and the Rajiv Gandhi Centre for Biotechnology, also have their own in-house fellowship programs for postdocs.

Growing Talent Early

The new funding may prove fruitful for innovation, but there is a need for greater access to education for Indians—approximately half of whom are under 25 years old. The government has responded to calls for greater educational opportunities for young people from a wider spectrum of society, and recently the DST launched the Innovation in Science Pursuit for Inspired Research (INSPIRE) program, with the aim of attracting students to science and expects to have funded one million young scholars by 2014.

This is just one program that seeks to build research capacity by giving students the opportunity to gain vital research-related skills. This is important in order to sustain global competitiveness among progressive nations like China. China, states Rao, currently produces almost as many journal articles as the United States, and he believes that the country will soon overtake the U.S. He further estimates that China graduates some 20,000 Ph.D.s annually. Says Rao, "How can we compete with this?"

Policymakers want India to increase the number of top scientific publications. "To achieve this we need more high-quality submissions, and to achieve that we need more good people," says Rao.

One of the challenges to finding "good people" is that many Indian students prefer to major in engineering rather than science, because of the promise and prestige of lucrative industrial career opportunities. But India's leaders recognized the need to motivate more youngsters to pursue science careers and hone research skills by forming five Indian Institutes of Science Education and Research (IISERs) in 2007. Here, faculty members have the freedom to pursue interdisciplinary projects while engaging their undergraduates in research.

"I watched K. Ganesh [the director of IISER Pune] build IISERs from the ground up," says Aseem Ansari, a professor of biochemistry at the **University of Wisconsin-Madison** (UW) in the United States and director of the Khorana Program, a cross-cultural exchange program for Indian and American students. "I believe IISERs are going to do for science (not just Indian science) what IITs did for technology and engineering. The first batch of students graduated recently and the impact of these 'research-oriented' students will be felt in the next five to 10 years."

The Khorana Program is an international consortium also designed to enhance research capacity within India and across borders. Jointly supported by UW, DBT, and the Indo-U.S. Science and Technology Forum (IUSSTF), and launched in 2008, the program grants Indian and American students the opportunity to pursue research at universities in each other's nations. continued>

IEEE Engineering in Medicine and Biology Society

Advancing Technology for Humanity

Join our Community!

Who are we?

The IEEE Engineering in Medicine and Biology Society (EMBS) is the world's largest international society of biomedical engineers. EMBS provides its members with access to the people, practices, information, ideas and opinions that are shaping one of the fastest growing fields in science.

Our Mission

The IEEE Engineering in Medicine & Biology Society benefits both humanity and its members by-

- advancing medicine and biology through the application of engineering sciences and technology
- promoting the profession of biomedical engineering, fostering professional development and recognizing excellence
- presenting conferences that bring together scientists, engineers and physicians from multiple disciplines to disseminate knowledge and solve complex problems
- establishing technical standards
- providing global leadership for the profession

Enjoy the benefits of an EMB membership with the

"Get to Know Us" product.

EMBS Publications

- IEEE PULSE: A Magazine of the IEEE Engineering in Medicine and Biology Society
- Transactions on Biomedical Engineering
- Journal of Biomedical and Health Informatics (formerly T-ITB)
- Transactions on Neural Systems and Rehabilitation Engineering
- Transactions on Medical Imaging
- Transactions on NanoBioscience
- Transactions on Computational Biology and Bioinformatics
- Transactions on Biomedical Circuits and Systems
- Reviews on Biomedical Engineering
- IEEE Journal on Translational Engineering in Health & Medicine (Launching in 2013)

EMBS Conferences

- Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)
- IEEE EMBS Special Topic Conference on Neural Engineering (NER)
- International Symposium on Biomedical Imaging (ISBI)
- International Conference on Biomedical Robotics & Biomechatronics (BIOROB)
- International Conference on Rehabilitation Robotics (ICORR)
- Healthcare Innovation (HIC)
- Micro & Nanotechnology in Medicine (MNM)
- Grand Challenges Conference Series (GCBE)
- IEEE EMBS International Conference on Biomedical & Health Informatics (BHI)
- IEEE EMBS Point-Of-Care Healthcare Technologies (Held in India in January 2013)

How to Join

For **$55.00**** you can experience the **EMBS Community** by trying out an **EMBS membership** for one year!

You will receive the following:

- Digital Access to IEEE PULSE
- Access to EMBS Electronic Resource, a digital library of all EMBS sponsored and co-sponsored publications (see listing on side) and EMBS sponsored, co-sponsored and technically sponsored conference proceedings (over 70,000 documents)

Join our Community by going to:
http://www.ieee.org/go/create_web_account
Create your Web Account
Click on the cart icon in the top right corner
Enter code: EMBGETKNOW (all caps)

IEEE Engineering in Medicine and Biology Society • 445 Hoes Lane, • Piscataway, NJ 08854
Email: emb exec@ieee.org • Website: www.embs.org

**Current IEEE and IEEE EMBS members are not eligible for this product. The offering is good for one year and can not be renewed. You will be eligible to apply for EMBS membership after one year.

Arindam Ghosh

Atmospheric Sciences

Satyajit Mayor

"We have mechanisms to hire scientists from anywhere in the world, and in fact many foreigners, are working at NCBS."

Developing Research Areas

The new funding policy will advance India's prowess in a number of strategic industries, such as space, energy, and the life sciences as well as important research areas in physics, materials science, and atmospheric science. Planned missions to Mars and a neutrino observatory will receive financial support under the new framework.

But it is also the new policy's acknowledgement of the role of innovation in targeted technological industries that is contributing to a renewed excitement among India's scientists. Energy is one of those strategic sectors, and many of India's scientific leaders are leveraging the government's interest in it to enhance vital research programs. For example, as materials science plays a central role in developing innovative technologies for the growing energy market, scientists like Arindam Ghosh at IISc are advancing materials research in graphene for solar cells and nanoelectronics.

India's scientists are also working with their counterparts in the United States on a major new Indo-U.S. initiative called the Solar Energy Research Institute for India and the United States (SERIIUS), funded through the U.S.-India Partnership to Advance Clean Energy and administered by IUSSTF. "This is an important project for India," says IISc's Chattopadhyay, leader of the Indian team. "India needs every drop of power we can produce."

Affordable health care and medical devices not requiring external electrical sources of power for operation are high-priority projects being undertaken at the CeNSE at IISc. "India has the largest number of people with diabetes in the world," says Rudra Pratap, chairperson of the CeNSE. "We want to use nanoelectronics to produce cheap wireless biosensors to monitor this disease in people living in rural India."

The Institute of Stem Cell Biology and Regenerative Medicine (inStem), and Centre for Cellular and Molecular Platforms (C-CAMP) on the NCBS campus are at the forefront of life sciences and translational medical research. "We offer an excellent environment for engaging in fundamental interdisciplinary research and also for early translation," says Satyajit Mayor, dean of NCBS. "We have mechanisms to hire scientists from anywhere in the world, and in fact many foreigners are working at NCBS."

NCBS is committed to translational research and recently hired S. Ramaswamy, a biophysicist who had been considered for the New York University associate dean of research, as the director of C-CAMP. "My mission is to find industrial applications for research being conducted at NCBS," he explains. "We want to encourage our researchers to move ideas from discovery to innovation." To do so, NCSB has partnered with industry leaders to accelerate commercialization. Recent examples of successful projects include an inexpensive kit to test for HIV/AIDS; the licensing of this technology is being negotiated with global companies.

This emphasis on translational research and technology transfer is being amplified throughout the country. The Tata Innovation Fellowship, a highly competitive scheme instituted by the DBT, recognizes and rewards innovative and productive life science researchers. Its specific prominence is on interdisciplinary, translational research with a potential for technology commercialization.

Future Challenges

Both challenges and hope lie ahead for India. Other often neglected but important science-related issues to address include establishing university curricula to improve the ability of young students to communicate in English, especially technical writing; the introduction of coordinated proactive strategies by research institutes to improve the 'visibility' of their scientists; incentives and financial support for entrepreneurial scientists to set up companies to commercialize ideas; and changes in labor laws to enable universities to hire qualified scientists irrespective of nationality.

The general mood of scientists—both veterans and young returnees—in India is positive. "India needs to build many more 'innovation ecosystems' like NCBS where there is a lot of intellectual freedom and a drive to define the cutting edge at an international level," notes UW's Ansari. With many institutions now focusing on nurturing research talent in undergraduates, "I am optimistic that this will yield the 'innovation dividend.'"

And as Indian Prime Minister Manmohan Singh recently said in an interview with *Science*, "One has to be optimistic. (....), unless one is optimistic, one is overwhelmed by the dimension of the development task that we have to accomplish."

Adarsh Sandhu is a freelance science writer based in Tokyo, Japan.
Alaina G. Levine, a science writer based in Tucson, AZ, contributed to this article.

DOI: 10.1126/science.opms.r1300128

India Alliance Margdarshi Fellowships

An opportunity for established bio-medical scientists of any Nationality, residing in India or overseas,

Who
- are leaders in their area of scientific research
- have been running an independent laboratory for 10 years or more
- are willing to relocate to a non-profit research organization in India and nucleate a cutting edge research programme to pursue research in India

Large and flexible funding for consumables, equipment, staff, travel to meetings and collaborative visits

No age limit. Clinical researchers with substantial research experience do not require a PhD to apply

Sponsored Applications due by 15, May 2013.

wellcome trust DBT
INDIA ALLIANCE

Images courtesy Wellcome Images

www.wellcomedbt.org
margdarshi@wellcomedbt.org

INSTITUTE OF LIFE SCIENCES

(An Autonomous Institute of the Dept. of Biotechnology, Government of India)

NALCO SQUARE, BHUBANESWAR-751 023

Advt. No. 2/2013

FACULTY POSITIONS AVAILABLE IN INDIA

Institute of Life Sciences, Bhubaneswar is an emerging multidisciplinary institute engaged in advanced research in biology. The Institute invites applications from Indian citizens for core scientific positions. Candidates with Ph.D/M.D. Degree and Post-Doctoral experience desirous of pursuing a dedicated research career in the field of Infectious Disease using any or all of the following tools of biology – Cell Biology, Immunology, Genomics, Proteomics, Nanotechnology, Bioinformatics and Systems Biology may send their Curriculum Vitae, list of publications, three names of reference(s) and proposed plan of research work to the **Director, Institute of Life Sciences, Nalco Square, Bhubaneswar – 751 023, India. Email: ilsindia.recruitment@gmail.com, Fax: 0091-674-2300728**.

Please visit **www.ils.res.in** for details about the Institute.

SciGenom

Defining genomics in India

We are a next-generation genomics company that provides unparalleled R&D environment for those who want to define genomics in India. Since our publication of the first female Indian whole genome, our R&D labs have embarked on a number of exciting projects. Besides supporting our internal efforts, our state-of-the-art genomics labs support application of next-generation sequencing technologies to help accelerate research in India and abroad. We specialize in variation detection, whole genome and exome sequencing, transcriptomics (RNA-seq), Chip-seq, metagenomics and bioinformatics analysis. Our NABL accredited medical genomics laboratory now has made several single gene tests for human disease affordable in India, while preparing for the application of next-gen sequencing in the clinic.

Offices:
India: Kochi & Chennai
USA: Pleasanton, CA

Info@scigenom.com
www.scigenom.com

Come join us for the genomics ride. Candidates with MS, PhD or MD/PhD training in oncology, genomics, bioinformatics, protein expression and purification, molecular biology are encouraged to apply to *hr@scigenom.com*

A proud sponsor of NGS conferences in INDIA.
Next NGS conference in Nov 2013 at New Delhi, India.
More information at *www.scigenomconferences.com*

Announcing Exceptional Opportunities for Faculty Positions in Indian Universities

UGC-Faculty Recharge Programme

University Grants Commission of India invites applications from outstanding highly motivated individuals, with demonstrated flair for research and teaching, for **positions of Assistant Professor, Associate Professor and Professor** in areas of basic sciences, i.e. Physics, Mathematics, Chemistry, Biology and Engineering and Earth Sciences.

These positions have been created as part of a new initiative designed to augment faculty resources in Indian Universities. The intent is to induct exceptional candidates with notable record of research in interdisciplinary/frontier areas of science, who will spearhead internationally competitive programmes along with their teaching responsibilities.

Initial appointments at each level shall be for a period of 5 years, extendible through peer evaluation by successive 5-year terms. There is provision for elevation to next higher rank following end-of-term or mid-term appraisal of those occupying the position of Assistant/Associate Professor. The positions are tenable till the age of superannuation (presently at 65 years).

The University Grants Commission will synergize and coordinate with Institutions under its umbrella for appropriate placement of successful candidates.

The UGC-faculty will receive emoluments at par with those of the Central University appointees. In addition, facilities (such as residential accommodation, etc.), commensurate with those available to faculty of the host Institution, may be extended to them. Also, they will be entitled to competitive start up grant for research.

For expanded description of the Programme, Positions and Application Procedure, please visit website: **www.ugcfrp.ac.in**

Contact address:

UGC-Faculty Recharge Programme
Old CRS Building
Jawaharlal Nehru University
Aruna Asaf Ali Marg, New Delhi – 110 067
Email: encor@ugcfrp.ac.in

INDIAN INSTITUTE OF SCIENCE EDUCATION AND RESEARCH MOHALI

The Indian Institute of Science Education & Research (IISER) Mohali, India invites applications from outstanding candidates for positions at the level of Assistant Professor, Associate professor and Professor, in the broad disciplines of Biology, Chemistry, Mathematics, Physics, Earth & Environmental Sciences and Humanities & social sciences.

IISER Mohali is a premier institute that has been set up by the Government of India to carryout research in frontline areas of science and to provide quality education at the undergraduate and graduate level. Located adjacent to Chandigarh, one of the most livable cities of India, IISER Mohali is into its 6th year and it provides both a supportive and a very vibrant research environment.

Successful applicants would be expected to maintain a strong research program and also participate in the undergraduate and graduate teaching. Salary and benefits will be commensurate with experience and Government of India norms (visit **http://www.iisermohali.ac.in/faculty_openings.html**).

Applicants should submit a letter of interest, stating their present and future research objectives, curriculum vitae that includes a summary of research and teaching, and contact information for three referees. A downloadable form for application is available from IISER Mohali website. Applications can be mailed to **Dean Faculty, IISER Mohali, Sector 81, Knowledge City, P.O. Manauli, Punjab 140306, India**; or by email to **deanfaculty@iisermohali.ac.in**

SCHOOL OF MEDICINE
Department of Pediatrics
UNIVERSITY OF COLORADO ANSCHUTZ MEDICAL CAMPUS

Children's Hospital Colorado

University of Colorado School of Medicine

Chief, Section of Genetics
Department of Pediatrics
University of Colorado School of Medicine and Children's Hospital of Colorado

The Department of Pediatrics of the University of Colorado School of Medicine and Children's Hospital Colorado seeks outstanding candidates for the position of Chief, Section of Genetics. The Section is home to a large and comprehensive clinical group, an accomplished group of researchers studying a wide array of genetic disorders, and operates full-service biochemical and molecular clinical laboratories.

The Section is based on the all-new Anschutz Medical Campus in Aurora, CO. The Anschutz Medical Campus includes the University of Colorado School of Medicine educational and research facilities, Children's Hospital Colorado, and the University of Colorado Hospital, as well as the Biomedical Graduate Programs, the Colorado School of Public Health, the School of Pharmacy and Pharmaceutical Sciences, the School of Dentistry and the College of Nursing.

The successful candidate is expected to lead the diverse activities of the Section, and to lead and build outstanding independent research programs in the area of medical genetics or genomics. Applicants must have a strong record of accomplishment in biomedical research, current NIH or other extramural funding, and a demonstrated commitment to education of medical and graduate students, fellows, and residents. M.D. appointees may be engaged in clinical care. Appointment will be at the Associate Professor or Professor level, as appropriate.

Applications are accepted electronically at **www.jobsatcu.com**, reference **Position F00057**.

TEXAS BIOMEDICAL RESEARCH INSTITUTE
SNPRC Southwest National Primate Research Center

STEM CELL REGENERATIVE MEDICINE FACULTY POSITIONS
Southwest National Primate Research Center (SNPRC)

The SNPRC at the Texas Biomedical Research Institute invites applications and nominations for two faculty positions in stem cell regenerative medicine. Applicants and nominees are expected to have an interest in, and preferably experience with, nonhuman primate models. Major strengths of the SNPRC are in genetics and genomics, metabolic disorders, and infectious diseases. Several primate models of human diseases developed at SNPRC are ideally suited for translational research on stem cell regenerative medicine. The SNPRC has an outstanding diversity of primate resources, including large colonies of baboons, rhesus monkeys, common marmosets, and chimpanzees.

The SNPRC has close associations with the University of Texas Health Science Center at San Antonio and the University of Texas at San Antonio, including opportunities for roles in graduate education. A critical mass of stem cell regenerative medicine researchers exists at those institutions, together with the developing stem cell regenerative medicine program at the SNPRC. The SNPRC has a strong tradition in postdoctoral training.

Rank and salary will be nationally competitive and commensurate with experience. The laboratories and offices of the successful candidates will be located in the new SNPRC laboratory and administration facility, which is now under construction and expected to be ready for occupancy in March, 2014. The positions, as well as temporary laboratory and office space, are available immediately.

Applications and nominations should be sent to the **Chair, SNPRC Regenerative Medicine Search Committee, c/o Human Resources Office, P.O. Box 760549, San Antonio, TX 78245-0549**, and should include a letter outlining qualifications and research interests. Applications, but not nominations, must also include a CV, and the names and contact information for at least three references. Additional information about the SNPRC can be found at **www.snprc.org**. Additional information about Texas Biomed can be found at **www.txbiomed.org**.

EOE

Ocean Research center of Zhoushan, Zhejiang University

浙江大学舟山海洋研究中心

Ocean Research Center of Zhoushan, Zhejiang University is a research institution in Zhoushan City, Zhejiang Province, China, which is set up by Zhoushan government and Zhejiang University in 2009. Multiple positions for senior talents are open in following areas:

- Marine sciences, including physical oceanography, marine chemistry, marine biology, marine geology etc.
- Marine engineering, including shipbuilding, mechatronics, environment, information, energy etc.
- Marine policy and culture.

Applicants should have strong research profile and potential capacity to conduct innovative research in these areas and have at least 3 years working experiences in first-class universities, research institutes or enterprises. The center provides state-of-the-art research facilities and strong supporting staffs. Competitive start-up support, salary and benefits will be offered according to individual qualification and experience.

Please submit your full CV, the cover letter and your future plan to **Ji Wenxiu** at **ajiwenxiu@yahoo.com.cn** or call **86-(0580)-2186309**. The positions will be open until they filled by appropriate candidates.

Icahn School of Medicine at Mount Sinai is one of the world's leading biomedical institutions and internationally acclaimed for excellence in scientific research, clinical care and education. It is among the nation's top twenty medical schools in NIH funding and U.S. News and World Report rankings. The School offers education programs leading to M.D., Ph.D. and master's degrees, and attracts outstanding students to its highly competitive programs and invigorating academic environment.

Icahn School of Medicine at Mount Sinai

CHAIR, DEPARTMENT OF PHARMACOLOGY AND SYSTEMS THERAPEUTICS

Mount Sinai's Department of Pharmacology and Systems Therapeutics is a vibrant, interactive department dedicated to research, teaching and development of novel therapeutics. The robust research portfolio of the Department places it among the top-ranked in the nation for NIH funding. Department faculty plays a critical role in the education of both graduate students and medical students. The Department has 30 full-time faculty members, and is one of eight basic science departments at Mount Sinai. Our leadership is committed to provide the resources needed not only to maintain but also to grow our outstanding Department of Pharmacology and Systems Therapeutics.

We seek a recognized leader with an outstanding academic background, including strong research credentials, commitment to education, mentoring experience and proven leadership and management skills. Direct experience in industry through biotechnology company formation or pharmaceutical employment and/or development of novel therapies is highly desirable, as are industry connections that will facilitate and foster translational research within the Department.

We offer a highly competitive remunerative package commensurate with the scope of this position, along with an opportunity to join our world class organization. Please email a letter of application or nomination, a curriculum vitae and three letters of reference to: **pharmsearch@mssm.edu**

To receive full consideration, applications should arrive by April 1, 2013.

Mount Sinai Medical Center is an equal opportunity/affirmative action employer. We recognize the power and importance of a diverse employee population and strongly encourage applicants with various experiences and backgrounds.

MSD China
Postdoctoral Research Fellow Program

At MSD, known as Merck in the United States, we work to bring important medicines and vaccines to people around the world through innovative science. In 2011, we established the MSD China Research and Development Center in Beijing. We are working with China's leading hospitals and institutes on research that will change the way healthcare is delivered to the Chinese people.

Building a Generation of Researchers

MSD China R&D is proud to announce the new MSD China Postdoctoral Research Fellow Program, an expansion of our recently launched MRL Postdoctoral Research Fellow Program that builds on our legacy of scientific excellence. The MSD China Program offers researchers in China a new path to the forefront of research in drug discovery.

As a Research Fellow in the MSD China Postdoctoral program, you will conduct research in important areas of drug discovery, either with scientists in the MSD China Research and Development Center, or in your current research institute in China under the tutelage of your postdoctoral advisor and a MSD scientist. Our postdocs will have the opportunity to:

- *Obtain unique experience in pharmaceutical research and development*
- *Work with leading researchers at Merck/MSD*
- *Generate innovative science resulting in high-quality external publications*
- *Present seminars at scientific meetings and have the opportunity to interact with the international scientific community*
- *Be positioned for choice careers in pharma, biotechnology, or academia*

MSD is a world's leading pharmaceutical research organization and has a long history of science-driven inventions that have improved human health. Our passion is improving health. This is what keeps us at the leading edge of scientific discovery and innovation.

To learn more about the MSD China Postdoctoral Research Fellow Program, or to apply for a position, visit http://

MSD 健康是福

CORP-1069238-0000

HUMAN FRONTIER SCIENCE PROGRAM (HFSP)

CALL FOR NOMINATIONS FOR THE 2014 HFSP NAKASONE AWARD

In keeping with its mission to promote innovative international research, HFSP invites nominations for the 2014 Nakasone Award which honors ground-breaking contributions in the life sciences. Typically these will be breakthroughs in understanding the complex mechanisms of living organisms that have important consequences for scientists throughout the world. Experimental, conceptual and technological contributions are eligible. This award recognizes the vision of former Prime Minister Nakasone of Japan in the creation of HFSP.

The winner of the 2013 award was Stephen Quake of Stanford University, USA, for pioneering work that advanced biological measuring techniques.

The competition is open; it is not limited to HFSP awardees and there is no age limit for candidates. However the jury will pay particular attention to recent breakthroughs by younger scientists. Nominations should be made before **5 April 2013** by submitting the standard one-page nomination form and the nominee's CV (see the HFSP website for more information). The selection will be made by the HFSP Council of Scientists at its meeting in July 2013.

The awardee will receive an unrestricted research grant of 10.000 USD, a commemorative medal and an invitation to deliver the Nakasone lecture at the 2014 HFSP Awardees Meeting.

HFSP, 12 quai Saint-Jean, 67080 STRASBOURG Cedex, France,
www.hfsp.org/awardees

Model Animal Research Center of Nanjing University
National Resource Center for Mutant Mice

南京大学模式动物研究所

Full-time Associate Professor or Full Professor

The Model Animal Research Center (MARC) of Nanjing University invites applications for full-time faculty positions, at academic ranks of Associate or Full Professor. Individuals with demonstrated accomplishments in, but not limited to, the following areas are encouraged to apply – neurobiology, immunology, development biology, metabolic disease and cancer biology. Highly competitive research support will be provided in an interactive and nurturing environment. Individuals will have an opportunity to establish a state-of-the-art independent research program in newly renovated space and to interact with a strong group of affiliated scientists. For more information and to apply visit website: **http://www.nicemice.cn/**.

Our institute's research interests emphasize on programs in the studies on human diseases using different model animal systems such as mouse, zebrafish, fruit fly, and nematode worm. Now MARC has established an excellent platform for functional analysis of transgenic and knockout mice, with AAALAC accredited SPF animal facility with 90,000 mouse cages till end of 2013. Interested individuals, regardless of their nationalities, should submit a detailed letter of interest, curriculum vitae, PDFs of three of their best publications, and three letters of recommendation to: **Dr. PAN Dejing, Deputy director of Model Animal Research Center, Nanjing University, 12 Xuefu Road, Nanjing, Jiangsu 210061, China** or preferably electronically to **pandj@nicemice.cn** or **pandj@nbri-nju.com**. The positions are available immediately. Applications will be evaluated by faculty search committee upon receipt until the positions are filled.

---

human development in landscapes
GRADUATE SCHOOL AT THE UNIVERSITY OF KIEL

C|A|U Christian-Albrechts-Universität zu Kiel

Christian-Albrechts-Universität zu Kiel, Germany aims to attract more qualified women for professorships. The Medical Faculty of Christian-Albrechts-Universität zu Kiel invites applications for a full-time

Junior Professor (W 1) position for Ancient DNA Analysis

initially for 3 years with a tenure track option starting as soon as possible. The Junior Professor will be employed as a civil servant (Beamtenverhältnis) on a temporary basis.

We are seeking an energetic and highly motivated scientist to conduct world-class research in the field of ancient DNA (aDNA) analysis. The applicant is expected to address research questions in Evolutionary Medicine and Archaeology with a focus on the genetic and genomic analysis of aDNA. Special emphasis will be placed on the establishment of a state-of-the-art aDNA laboratory and the development of innovative research methods, e.g. Next Generation Sequencing technologies.

A PhD in a discipline relevant for the project and qualifications in both molecular life science and pre- and protohistory are a prerequisite. Preference will be given to individuals with a solid publication record.

The successful candidate possesses excellent knowledge in the field of genome sequencing and data analysis as well as broad experience in aDNA lab work, standard molecular genetic methods, archaeological theories and archaeological fieldwork. Familiarity with setting up and managing an aDNA laboratory is advantageous.

The position will be assigned to the Johanna-Mestorf-Academy, and a major aspect of this professorship is the contribution to the interdisciplinary research environment of this academy. Close collaborations with institutes and scientists from various disciplines is explicitly expected. The professor will be involved in the supervision of PhD students and teaching of undergraduates. The position is primarily based in the Medical Faculty and has secondary memberships in the Philosophical and Mathematical-Natural Sciences Faculties.

Applicants must have the necessary formal qualifications as set out in § 64 of the Universities and Colleges Act of Schleswig-Holstein (HSG). For more information, please refer to the webpage www.berufungen.uni-kiel.de/de. After a positive evaluation and if the qualifying requirements according to § 62 HSG are met, a tenure option will be offered, i.e. a change of status into a permanent W 2-professorship of the position. At Christian-Albrechts-Universität zu Kiel, for this a separate evaluation process will be performed during the second phase of the junior professorship in addition to the standard evaluation procedure for junior professors.

The HSG asks the Medical Faculties of Christian-Albrechts-Universität zu Kiel and Universität zu Lübeck to collaborate closely with each other and with Universitätsklinikum Schleswig-Holstein, to establish and coordinate research foci. Moreover, the federal state of Schleswig-Holstein expects clinics, institutes and newly appointed professors to collaborate with each other.

Women with equivalent qualifications, competence and expertise will be given preference. The university therefore strongly encourages women with appropriate qualifications to apply for the position. The university supports the employment of disabled persons. Persons with disabilities will, with appropriate qualifications and aptitudes, be employed preferentially.

Applications, including a curriculum vitae, qualifying documentation (publications, evidence of external funding, teaching experience) and a short research plan should be addressed to the Dean of the Medical Faculty of Kiel University, Christian-Albrechts-Universität zu Kiel, Christian-Albrechts-Platz 4, 24098 Kiel, Germany.

The closing date is **29th of March 2013.** Candidates willing to apply should read the application guideline and fill in the profile form available on our website:
www.uni-kiel.de/landscapes/allgemein/jobs/jrprof_adna.shtml

AAAS is here – helping scientists achieve career success.

Every month, over 400,000 students and scientists visit ScienceCareers.org in search of the information, advice, and opportunities they need to take the next step in their careers.

A complete career resource, free to the public, *Science* Careers offers a suite of tools and services developed specifically for scientists. With hundreds of career development articles, webinars and downloadable booklets filled with practical advice, a community forum providing answers to career questions, and thousands of job listings in academia, government, and industry, *Science* Careers has helped countless individuals prepare themselves for successful careers.

As a AAAS member, your dues help AAAS make this service freely available to the scientific community. If you're not a member, join us. Together we can make a difference.

To learn more, visit aaas.org/plusyou/sciencecareers

AAAS + U = Δ

AAAS is here – promoting universal science literacy.

In 1985, AAAS founded Project 2061 with the goal of helping all Americans become literate in science, mathematics, and technology. With its landmark publications *Science for All Americans* and *Benchmarks for Science Literacy*, Project 2061 set out recommendations for what all students should know and be able to do in science, mathematics, and technology by the time they graduate from high school. Today, many of the state standards in the United States have drawn their content from Project 2061.

Every day Project 2061 staff use their expertise as teachers, researchers, and scientists to evaluate textbooks and assessments, create conceptual strand maps for educators, produce groundbreaking research and innovative books, CD-ROMs, and professional development workshops for educators, all in the service of achieving our goal of universal science literacy.

As a AAAS member, your dues help support Project 2061 as it works to improve science education. If you are not yet a AAAS member, join us. Together we can make a difference.

To learn more, visit **aaas.org/plusyou/project2061**

AAAS + U = Δ

AAAS is here – bringing scientific expertise to policy making.

Good science policy is the result of politicians understanding science and scientists understanding policy. Toward this end, AAAS manages the Science & Technology Policy Fellowships program, which embeds scientists and engineers in the federal government for up to two years. From Congress to the State Department, each class of Fellows contributes to the policy-making process while getting hands-on experience at the intersection of science and policy. As a AAAS member your dues support these efforts. If you're not yet a AAAS member, join us. Together we can make a difference.

To learn more, visit **aaas.org/plusyou/fellows**

AAAS + U = Δ

AAAS is here – Science Funding, Climate Regulation, Human Rights.

Around the world, governments turn to AAAS as an objective, multidisciplinary scientific authority to educate public officials and judicial figures on today's most pressing issues. And this is just one of the ways that AAAS is committed to advancing science to support a healthy and prosperous world. Join us. Together we can make a difference.

To learn more, visit **aaas.org/plusyou/policy**

AAAS + U = Δ

POSITIONS OPEN

CAREER OPPORTUNITY—Doctor of Optometry (OD) degree in 27 months for Ph.D.s in science and M.D.s. Excellent career opportunities for O.D.-Ph.D.s and O.D.-M.D.s in research, education, industry and clinical practice. This unique program starts in March of each year, features small classes, and 12 months devoted to clinical care.

Contact the Admissions Office, **telephone: 800-824-5526** at **The New England College of Optometry, 424 Beacon Street, Boston, MA 02115.** Additional information, **website: http://www.neco.edu; e-mail: admissions@neco.edu.**

MARKETPLACE

Stem Cell Lawsuit Finally Over. Russian Team Retrieves First Sample from Lake Vostok. Surprise Choices Mark New Leadership on U.S. House Science Panel. India Unveils Ambitious Science Policy.

Now, more than ever, developments in the lab are directly connected to decisions made in the halls of government.

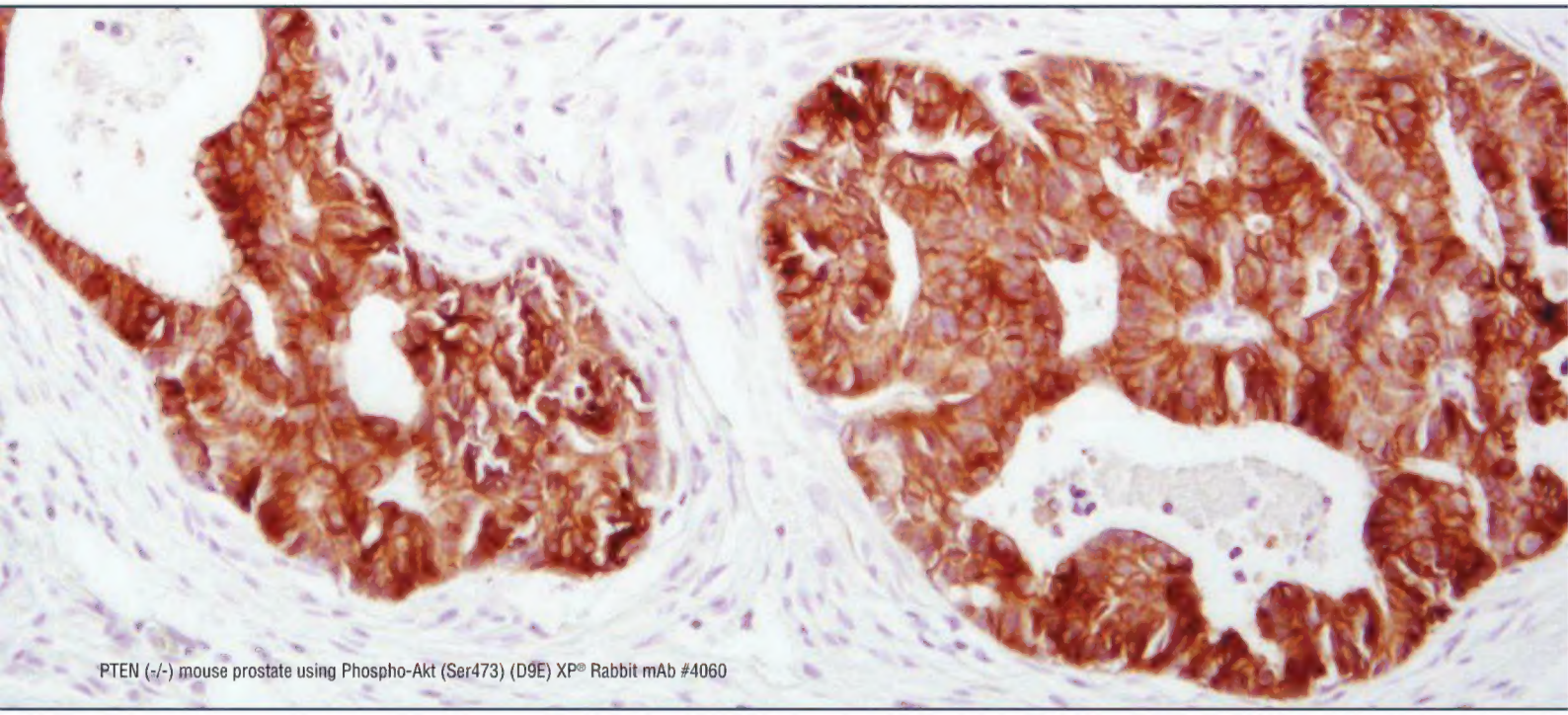
*Science*Insider, the policy blog from the journal *Science*, is your source for news from the intersection of science and policy. From budget debates in the United States Congress, to climate change agreements at the United Nations, *Science*Insider covers the issues that have an impact on your work, your field, and your world.

Keep up to date and keep informed. Go inside the issues at **www.ScienceInsider.org**

*Science*Insider

Breaking news and analysis from the world of science policy

Science
AAAS



PTEN (-/-) mouse prostate using Phospho-Akt (Ser473) (D9E) XP® Rabbit mAb #4060

XP® Monoclonal Antibodies,
one antibody, multiple applications

Unparalleled product quality, validation, and technical support.

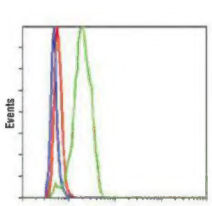
XP® monoclonal antibodies are a line of high quality rabbit monoclonal antibodies exclusively available from Cell Signaling Technology. Any product labeled with XP has been carefully selected based on superior performance in the most relevant research applications.

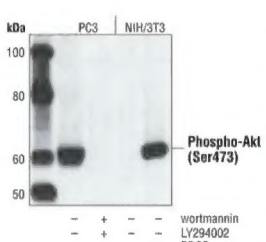
Exceptional specificity
As with all of our antibodies, the antibody is specific to your target of interest, saving you valuable time and resources.

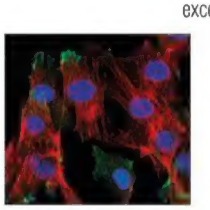
+ Exceptional sensitivity
The antibody will provide a stronger signal for your target protein in cells and tissues, allowing you to monitor expression of low levels of endogenous proteins, saving you valuable materials.

+ Exceptional stability and reproducibility
XMT technology combined with our stringent quality control ensures maximum lot-to-lot consistency and the most reproducible results.

= Exceptional performance™
XMT technology coupled with our extensive antibody validation and stringent quality control delivers XP monoclonal antibodies with exceptional performance in the widest range of applications.






Insulin-treated C2C12 cells

◀ ***Phospho-Akt (Ser473) (D9E) XP® Rabbit mAb #4060 demonstrates exceptional performance in a wide range of applications.***

For experimental details, additional information, and a complete list of available XP® monoclonal antibodies visit…
www.cellsignal.com/sci

Cell Signaling TECHNOLOGY®

Orders (toll-free) 1-877-616-2355 | **Technical support** (toll-free) 1-877-678-8324 support@cellsignal.com | **Inquiries** info@cellsignal.com | **Environmental Commitment** eco.cellsignal.com